16864N-093500

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application: ２００３年　３月　７日

出願番号
Application Number: 特願２００３－０６１４０３

[ST. 10/C]: [ＪＰ２００３－０６１４０３]

出願人
Applicant(s): 株式会社日立製作所

２００３年　７月３１日

特許庁長官
Commissioner,
Japan Patent Office 今井康夫

出証番号　出証特２００３－３０６１２６５

【書類名】　特許願
【整理番号】　NT02P0644
【提出日】　平成15年 3月 7日
【あて先】　特許庁長官　殿
【国際特許分類】　G06F　3/06
【発明者】
【住所又は居所】　神奈川県川崎市麻生区王禅寺１０９９番地　株式会社日立製作所　システム開発研究所内
【氏名】　田中　勝也
【発明者】
【住所又は居所】　神奈川県川崎市麻生区王禅寺１０９９番地　株式会社日立製作所　システム開発研究所内
【氏名】　上村　哲也
【特許出願人】
【識別番号】　000005108
【氏名又は名称】　株式会社日立製作所
【代理人】
【識別番号】　100068504
【弁理士】
【氏名又は名称】　小川　勝男
【電話番号】　03-3661-0071
【選任した代理人】
【識別番号】　100086656
【弁理士】
【氏名又は名称】　田中　恭助
【電話番号】　03-3661-0071

【選任した代理人】

【識別番号】　100094352

【弁理士】

【氏名又は名称】　佐々木　孝

【電話番号】　03-3661-0071

【手数料の表示】

【予納台帳番号】　081423

【納付金額】　21,000円

【提出物件の目録】

【物件名】　明細書　1

【物件名】　図面　1

【物件名】　要約書　1

【プルーフの要否】　要

【書類名】　明細書

【発明の名称】　ディスクアレイ装置および障害回復制御方法

【特許請求の範囲】

【請求項１】

チャネルパスを介して上位装置に接続されたディスクコントローラと、上記ディスクコントローラに接続された保守端末と、ディスクチャネルを介して上記ディスクコントローラに接続されたディスクアレイとからなり、

上記ディスクアレイが、それぞれ基板内配線に接続された複数のディスクドライブを搭載した複数のドライブ基板と、上記各ドライブ基板の基板内配線をディスクチャネル用の配線と接続するための複数のコネクタを備えた共通基板とからなり、

上記複数のドライブ基板が上記複数のコネクタを介して上記共通基板に着脱可能に搭載され、上記ディスクコントローラが、上記ディスクアレイ内のドライブ基板の一部を予備系、残りを現用系と定義し、現用系のドライブ基板群において互いに異なるドライブ基板上に搭載されたＮ＋１個（Ｎ≧２）のディスクドライブを論理グループとして管理し、各論理グループ内で生成された誤り訂正情報の格納領域を上記複数のディスクドライブに分散的または特定ディスクドライブに固定的に割当てて、上記ディスクアレイのデータの書き込みと読み出しを制御するようにしたディスクアレイ装置において、

上記ディスクコントローラが、上記ディスクアレイ内の何れかの現用系ディスクドライブが障害ドライブとなった時、上記障害ドライブが搭載された障害基板上の各ディスクドライブの格納データと同一のデータを上記予備系のドライブ基板群の中から選択された交替基板上のディスクドライブに蓄積した後、上記障害基板上の各ディスクドライブが所属する論理グループを上記交替基板上のディスクドライブを含む新たな構成に再編成するための手段と、

上記論理グループ再編成の完了後に、上記保守端末に対して上記障害基板が交換可能状態となったこと通知するための手段を備えたことを特徴とするディスクアレイ装置。

【請求項２】

前記論理グループの再編成手段が、前記障害ドライブに代わる前記交替基板上の１つの交替ディスクドライブに対して、上記障害ドライブと同一論理グループに所属した他の複数のディスクドライブからの読み出しデータに基づいて再生されたデータを書き込むための手段と、前記障害基板上に搭載された正常ディスクドライブからの読み出しデータを前記交替基板上の別の交替ディスクドライブに順次に複写するための手段とを含むことを特徴とする請求項１に記載のディスクアレイ装置。

【請求項３】

前記共通基板が、前記ディスクチャネル用の配線から前記ドライブ基板接続用の各コネクタを選択的にバイパスするための複数のバイパス回路を備え、

前記ディスクコントローラが、前記論理グループ再編成の完了後に、前記障害基板の接続コネクタと対応するバイパス回路をバイパス状態に切替えるための手段を備えたことを特徴とする請求項１または請求項２に記載のディスクアレイ装置。

【請求項４】

前記ディスクコントローラが、前記障害基板の接続コネクタに復旧基板が再接続された時、該復旧基板を予備系のドライブ基板として管理することを特徴とする請求項１または請求項２に記載のディスクアレイ装置。

【請求項５】

前記ディスクコントローラが、前記ディスクアレイを構成するドライブ基板を正常状態、交換待ち状態、予備状態の順に遷移する状態コードによって管理するための基板管理テーブルを備え、状態コードが正常状態の基板を前記現用系、予備状態の基板を前記予備系とすることを特徴とする請求項１または請求項２に記載のディスクアレイ装置。

【請求項６】

前記共通基板上で前記ドライブ基板接続用のコネクタがＸ、Ｙ軸上の座標値で特定される２次元配置を有し、各ドライブ基板上の複数のディスクドライブがＺ軸方向に配列され、

前記ディスクコントローラが、互いに同一のＸ座標値、Ｚ座標値を有し、Ｙ座

標値が異なるＮ＋１個のディスクドライブによって前記論理グループを形成し、

前記論理グループ再編成手段が、前記予備系のドライブ基板群の中から前記障害基板と同一のＹ座標値をもつドライブ基板を前記交替基板として選択し、上記故障基板上の複数のディスクドライブと上記交替基板上の交替ディスクドライブとをそれぞれのＺ座標値に従って対応付けることを特徴とする請求項１または請求項２に記載のディスクアレイ装置。

【請求項７】

上位装置と保守端末に接続されたディスクコントローラと、ディスクチャネルを介して上記ディスクコントローラに接続されたディスクアレイとからなり、

上記ディスクアレイが、それぞれ基板内配線に接続された複数のディスクドライブを搭載した複数のドライブ基板と、上記各ドライブ基板の基板内配線をディスクチャネル用の配線と接続するための複数のコネクタを備えた共通基板とからなり、

上記複数のドライブ基板が上記複数のコネクタを介して上記共通基板に着脱可能に搭載され、上記ディスクコントローラが、上記ディスクアレイ内のドライブ基板の一部を予備系、残りを現用系と定義し、現用系のドライブ基板群において互いに異なるドライブ基板上に搭載されたＮ＋１個（Ｎ≧２）のディスクドライブを論理グループとして管理し、各論理グループ内で生成された誤り訂正情報の格納領域を上記複数のディスクドライブに分散的または特定ディスクドライブに固定的に割当てて、上記ディスクアレイのデータの書き込みと読み出しを制御するようにしたディスクアレイ装置における上記ディスクコントローラによる障害回復制御方法であって、

上記ディスクアレイ内の何れかの現用系ディスクドライブが障害ドライブとなった時、上記予備系のドライブ基板群の中から、上記障害ドライブが搭載された障害基板に代わる交替基板を選択するステップと、

上記障害ドライブが搭載された障害基板上の各ディスクドライブの格納データと同一のデータを上記予備系のドライブ基板群の中から選択された交替基板上のディスクドライブに蓄積するステップと、

上記障害基板上の各ディスクドライブが所属する論理グループを上記交替基板

上のディスクドライブを含む新たな構成に再編成するステップと、

上記論理グループの再編成後に、上記保守端末に対して上記障害基板が交換可能状態となったこと通知するステップとを含むことを特徴とする障害回復制御方法。

【請求項８】

前記交替基板上のディスクドライブへのデータ蓄積ステップが、

前記障害ドライブに代わる前記交替基板上の１つの交替ディスクドライブに対して、上記障害ドライブと同一論理グループに所属した他の複数のディスクドライブからの読み出しデータに基づいて再生されたデータを書き込むステップと、

前記障害基板上に搭載された正常ディスクドライブからの読み出しデータを前記交替基板上の別の交替ディスクドライブに順次に複写するステップとからなることを特徴とする請求項７に記載の障害回復制御方法。

【請求項９】

前記共通基板が、前記ディスクチャネル用の配線から前記ドライブ基板接続用の各コネクタを選択的にバイパスするための複数のバイパス回路を備え、

前記ディスクコントローラが、前記論理グループの再編成後に、前記障害基板の接続コネクタと対応するバイパス回路をバイパス状態に切替えるステップを含むことを特徴とする請求項７または請求項８に記載の障害回復制御方法。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

本発明は、ディスクアレイ装置および障害回復制御方法に関し、更に詳しくは、冗長ディスクドライブの含む複数のディスクドライブで１つの論理グループを形成してディスク障害に備えるＲＡＩＤ形式のディスクアレイ装置および障害回復制御方法に関する。

【０００２】

【従来の技術】

高性能のコンピュータシステムは、大容量の２次記憶装置を備え、ＣＰＵ等の上位装置が必要とするデータを上記２次記憶装置に随時にリード・ライトしてい

る。２次記憶装置としては、例えば、磁気ディスクや光ディスクのように、ランダムアクセス可能な不揮発性記憶媒体をもつディスク装置が一般的であり、最近では、記憶容量を増加するために多数の小型ディスクドライブ（以下、単にドライブと言う）からなるディスクアレイ装置が主流になってきている。

【０００３】

ディスクアレイ装置では、少なくとも1個の冗長ドライブを含む複数のドライブで1つの論理グループを形成し、論理グループ毎にドライブ障害に対処できるようにしたＲＡＩＤ（Redundant Array of Inexpensive Disk）方式が採用されている。

【０００４】

ＲＡＩＤには標準化された幾つかのレベルがあり、例えば、ＲＡＩＤ１（レベル１）のディスクアレイ装置では、データ格納用の各ドライブと対をなして冗長ドライブ（予備ドライブ）を用意しておき、同一データを２つのドライブに並列的に書き込むことによって、何れかのデータドライブに障害が発生したとき、これと対をなす予備ドライブから必要なデータを読み出せるようにしている。

【０００５】

ＲＡＩＤ３（レベル３）では、Ｎ＋１個（Ｎ≧２）のドライブで１つの論理グループ（パリティグループまたはＲＡＩＤグループと言う場合もある）を形成し、そのうちの１個を誤り訂正情報（以下、パリティで代表させる）格納用の冗長ドライブ、残りをデータ格納用ドライブとしている。本明細書では、誤り訂正情報をパリティで代表させるが、各論理グループで生成される誤り訂正情報にはパリティ以外の他の情報を適用できること明らかである。

【０００６】

ＲＡＩＤ３のディスクアレイ装置では、上位装置からデータブロックの書き込み要求があった時、書き込み要求に付随するデータブロックを固定長（例えば、1バイト長）の複数のサブデータブロックに分割し、これらのサブデータブロックを上記Ｎ個のデータ格納用ドライブに順次に振り分けて格納する。冗長ドライブには、データ格納用ドライブ内で互いに同一のアドレスをもつ同一論理グループ内のＮ個のサブデータブロックから生成された誤り訂正情報（例えば、パリテ

ィデータ）が格納される。上位装置からデータブロックの読み出し要求があった時は、上記Ｎ個のデータ格納用ドライブから並列的にサブデータブロックを読み出し、これらのサブデータブロックを所定の順序で結合することによって、元のデータブロックを復元している。

【０００７】

ＲＡＩＤ４（レベル４）でも、Ｎ＋１個（Ｎ≧２）のドライブで１つの論理グループを形成し、そのうちの１個を誤り訂正情報格納用の冗長ドライブ、残りをデータ格納用ドライブとしている。但し、ＲＡＩＤ４のディスクアレイ装置では、上位装置からデータブロックの書き込み要求があった時、１つの書き込み要求に付随するデータブロックを上記何れかのデータ格納用ドライブに格納し、次の書き込み要求に付随するデータブロックは別のデータ格納用ドライブに格納する形式で、データの書き込みが行われる。従って、冗長ドライブには、データ格納用ドライブ内で同一のアドレスをもつそれぞれが別のデータブロックに所属するデータ部分から生成された誤り訂正情報が格納されることになる。

【０００８】

ＲＡＩＤ５（レベル５）では、レベル４と同様に、１つの書き込み要求に付随するデータブロックでのデータ書き込みが行なわれる。但し、誤り訂正情報の格納領域は、レベル４にように特定のディスクドライブに固定的に割当てられるのではなく、論理グループを形成する複数（Ｎ＋１）のディスクドライブに分散的に割当てられる。

【０００９】

ＲＡＩＤ３～ＲＡＩＤ５のディスクアレイ装置では、何れかのドライブに障害が発生した場合、障害ドライブと同一論理グループに所属する他のドライブから読み出したデータに基づいて、上記障害ドライブが保持していたデータまたは誤り訂正情報（例えば、パリティデータ）を再生することが可能となる。

【００１０】

上述したディスクアレイ装置で、記憶容量を大容量化し、筐体サイズを小型化するためには、狭い空間にできるだけ多数のドライブを実装する必要がある。一般的には、それぞれが制御用のＬＳＩチップを搭載した複数の制御基板と、それ

ぞれが複数のドライブを搭載した多数のドライブ基板とをマザーボード上に並列的に配設されたコネクタに差込み、各ドライブをドライブ基板上の配線を介してマザーボード上のディスクチャネル配線に結合するようにした装置構成が採用される。

【００１１】

この構成において、ドライブ障害が発生した時、ドライブ基板から障害ドライブだけを取り外して新たなドライブと交換しようとすると、隣接するドライブ基板の間にドライブ交換作業に必要な空間を残す必要があり、マザーボード上でのドライブ基板の実装密度が低下する。

【００１２】

この問題点に着目した従来技術として、例えば、特開平７－２３０３６２号公報（特許文献１）には、マザーボードに多数のドライブ基板を高密度で実装しておき、ドライブ障害が発生した時は、障害ドライブが搭載されているドライブ基板（障害基板）そのものをマザーボードから取り外し、装置筐体の外部において障害ドライブを正常ドライブに交換し、部品交換されたドライブ基板をマザーボードに再接続するようにしたディスクアレイ装置が提案されている。

【００１３】

上記構成によれば、部品交換期間中に、障害ドライブのみならず、障害基板上に搭載されている複数の正常ドライブもディスクアレイから除外されてしまう。

そこで、上記特許文献１では、それぞれ異なるドライブ基板上に搭載されているＮ＋１個のドライブによって各論理グループを形成しておき、上位装置から障害ドライブの格納データ、または障害基板の取り外しによって不在となったドライブの格納データに対して読み出し要求があった時は、障害ドライブまたは不在ドライブと同一論理グループに所属する他の複数のドライブからデータを読み出し、これらのデータによって要求データを再生するようにしている。

【００１４】

また、上記特許文献１では、障害基板の取り外し前に障害ドライブに対して書き込み要求があったデータと、障害基板の取り外し後に不在となったドライブ（障害ドライブと複数の正常ドライブ）に対して書き込み要求があったデータにつ

いては、ディスク制御装置が備えるキャッシュメモリに格納しておき、部品交換されたドライブ基板がマザーボードに再接続された時、キャッシュメモリから復旧基板上の該当ドライブにデータを書き移すことを提案している。尚、障害ドライブの除去によって読み出し不能となった損失データについては、ドライブ基板が再接続された時、障害ドライブと同一論理グループに所属する他の複数のドライブからの読み出しデータによって再生し、復旧基板上の交替ドライブに書き込むようにしている。

【００１５】

上記特許文献１では、更に、各論理グループと対応して予備ドライブを用意しておき、上位装置から不在ドライブに対して書き込み要求のあったデータをキャッシュメモリの代わりに上記予備ドライブに一時的に格納しておき、部品交換されたドライブ基板がマザーボードに再接続された時、上記予備ドライブから復旧基板上の該当ドライブにデータを書き移すことも提案している。また、障害ドライブに代わる予備ドライブの蓄積データが大量になった場合は、部品交換されたドライブ基板が再接続された時、予備ドライブ（または障害ドライブ）と同一論理グループに所属する他の複数のドライブからの読み出しデータによって障害ドライブの損失データを再生し、これらのデータを上記予備ドライブに書き込むことによって、予備ドライブを正常ドライブとして引き続き利用してもよい旨の提案もしている。

【００１６】

【特許文献１】

特開平７－２３０３６２号

【００１７】

【発明が解決しようとする課題】

然るに、特許文献１で提案されたディスクアレイ装置は、筐体サイズの小型化と記憶容量の大容量化に好適なハードウエア構成と言えるが、障害ボードの取り外し期間中に上位装置からデータの読み出し要求があった時、障害ドライブのみならず、不在となった複数の正常ドライブについても、論理グループを利用したデータの再生動作が必要となり、障害基板の取り外し期間中のデータ読み出し要

求に対して応答遅れが発生するという問題があった。

【００１８】

また、上記従来技術では、部品交換されたドライブ基板がマザーボードに復帰した時に、キャッシュメモリまたは予備ドライブから復旧ボード上のドライブへのデータ書き込みと、障害ドライブ損失データの再生動作とを実行しているため、ディスクアレイの正常状態への復帰が遅れるという問題があった。

【００１９】

本発明の目的は、筐体サイズの小型化と記憶容量の大容量化に適し、ドライブ障害が発生しても、速やかに正常状態に復帰できるディスクアレイ装置および障害回復制御方法を提供することにある。

本発明の他の目的は、ドライブ障害が発生した時、保守員による障害ドライブの部品交換に十分な作業時間を許容できるディスクアレイ装置および障害回復制御方法を提供することにある。

【００２０】

【課題を解決するための手段】

上記目的を達成するため、本発明のディスクアレイ装置および障害回復制御方法では、ドライブ障害が発生した時、障害ドライブを搭載した障害基板上の複数のディスクドライブについて、交替ディスクドライブへのデータの移し替えを行い、論理グループの再編成が完了した後に、障害基板の交換を行うようにしたことを特徴とする。

【００２１】

障害ドライブの格納データは、同一論理グループに所属した他のドライブからの読み出しデータに基づいて再生する必要があるが、障害基板上の正常ドライブの格納データは、各ドライブからの読み出しデータをそのまま交替ディスクドライブに複写すれば済むため、正常ドライブから交替ディスクドライブへのデータの移し替えは短時間で終了できる。本発明によれば、論理グループの再編成が完了した後に障害基板の交換を行うようにしているため、障害基板の取り外し期間中でも、通常のデータ・リード／ライトが可能となり、従来技術のように、不在ドライブに対する論理グループによるデータ再生を行う必要がない。

【００２２】

更に詳述すると、本発明は、チャネルパスを介して上位装置に接続されたディスクコントローラと、上記ディスクコントローラに接続された保守端末と、ディスクチャネルを介して上記ディスクコントローラに接続されたディスクアレイとからなり、上記ディスクアレイが、それぞれ基板内配線に接続された複数のディスクドライブを搭載した複数のドライブ基板と、上記各ドライブ基板の基板内配線をディスクチャネル用の配線と接続するための複数のコネクタを備えた共通基板とからなり、上記複数のドライブ基板が上記複数のコネクタを介して上記共通基板に着脱可能に搭載され、上記ディスクコントローラが、上記ディスクアレイ内のドライブ基板の一部を予備系、残りを現用系と定義し、現用系のドライブ基板群において互いに異なるドライブ基板上に搭載されたＮ＋１個（Ｎ≧２）のディスクドライブを論理グループとして管理し、各論理グループ内で生成された誤り訂正情報の格納領域を上記複数のディスクドライブに分散的または特定ディスクドライブに固定的に割当てて、上記ディスクアレイのデータの書き込みと読み出しを制御するようにしたディスクアレイ装置において、

上記ディスクコントローラが、上記ディスクアレイ内の何れかの現用系ディスクドライブが障害ドライブとなった時、上記障害ドライブが搭載された障害基板上の各ディスクドライブの格納データと同一のデータを上記予備系のドライブ基板群の中から選択された交替基板上のディスクドライブに蓄積した後、上記障害基板上の各ディスクドライブが所属する論理グループを上記交替基板上のディスクドライブを含む新たな構成に再編成するための手段と、

上記論理グループ再編成の完了後に、上記保守端末に対して上記障害基板が交換可能状態となったこと通知するための手段を備えたことを特徴とする。

【００２３】

本発明の１実施例によれば、上記論理グループの再編成手段が、障害ドライブに代わる交替基板上の１つの交替ディスクドライブに対して、上記障害ドライブと同一論理グループに所属した他の複数のディスクドライブからの読み出しデータに基づいて再生されたデータを書き込むための手段と、障害基板上に搭載された正常ディスクドライブからの読み出しデータを上記交替基板上の別の交替ディ

スクドライブに順次に複写するための手段とを含む。

【００２４】

本発明の他の特徴は、共通基板が、ディスクチャネル用の配線からドライブ基板接続用の各コネクタを選択的にバイパスするための複数のバイパス回路を備え、上記ディスクコントローラが、上記論理グループ再編成の完了後に、障害基板の接続コネクタと対応するバイパス回路をバイパス状態に切替えるための手段を備えることにある。

【００２５】

本発明の更に他の特徴は、上記ディスクコントローラが、ディスクアレイを構成するドライブ基板を正常状態、交換待ち状態、予備状態の順に遷移する状態コードによって管理するための基板管理テーブルを備え、状態コードが正常状態の基板を現用系、予備状態の基板を予備系として扱うことにある。

【００２６】

本発明の１実施例によれば、共通基板上で前記ドライブ基板接続用のコネクタがＸ、Ｙ軸上の座標値で特定される２次元配置を有し、各ドライブ基板上の複数のディスクドライブがＺ軸方向に配列され、上記ディスクコントローラが、互いに同一のＸ座標値、Ｚ座標値を有し、Ｙ座標値が異なるＮ＋１個のディスクドライブによって論理グループを形成する。また、前記論理グループ再編成手段が、予備系のドライブ基板群の中から故障基板と同一のＹ座標値をもつドライブ基板を交替基板として選択し、上記障害基板上の複数のディスクドライブと上記交替基板上の交替ディスクドライブとをそれぞれのＺ座標値に従って対応付ける。

【００２７】

本発明によるディスクアレイ装置における障害回復制御方法は、ディスクコントローラが、

ディスクアレイ内の何れかの現用系ディスクドライブが障害ドライブとなった時、予備系のドライブ基板群の中から、上記障害ドライブが搭載された障害基板に代わる交替基板を選択するステップと、

上記障害ドライブが搭載された障害基板上の各ディスクドライブの格納データと同一のデータを上記予備系のドライブ基板群の中から選択された交替基板上の

ディスクドライブに蓄積するステップと、

上記障害基板上の各ディスクドライブが所属する論理グループを上記交替基板上のディスクドライブを含む新たな構成に再編成するステップと、

上記論理グループの再編成後に、上記保守端末に対して上記障害基板が交換可能状態となったこと通知するステップとを含むことを特徴とする。

本発明のその他の特徴は、図面を参照した以下の実施例の説明から明らかになる。

【００２８】

【発明の実施の形態】

以下、本発明の実施例について、図面を参照して説明する。

図１は、本発明が適用されるディスクアレイ装置の1実施例を示す。

ここに示したディスクアレイ装置は、複数のディスクドライブからなるディスクアレイ５を２台のディスクコントローラ１Ａ、１Ｂからアクセスする冗長構成となっている。これらのディスクコントローラ１Ａ、１Ｂは、表示装置９を備えたサービスプロセッサ（ＳＶＰ）２に接続されている。

【００２９】

ディスクコントローラ１Ａは、複数本のチャネルパス３Ａ（３０Ａ～３３Ａ）を介して上位装置となるＣＰＵ（図示せず）に接続されるチャネルアダプタ１０と、後述するバイパス制御用の信号線を含む複数本のディスクチャネル４Ａ（４０Ａ～４３Ａ）を介してディスクアレイ５に接続されるディスクアダプタ（ＤＫＡ）２０と、これらの要素を相互に接続する相互結合網４０とからなる。

【００３０】

本実施例では、各論理グループ（以下、ＲＡＩＤグループと言う）を構成するディスクドライブ数を４個とし、各ディスクドライブのロジカルユニット数を１個として説明するが、本発明は、ＲＡＩＤグループのコンポーネントとなるドライブ個数と各ディスクドライブのロジカルユニット数を実施例に限定するものではない。

【００３１】

ディスクアレイ５は、複数のディスクドライブPDEV0～PDEV127からなり、これ

らのディスクドライブは、ディスクチャネルと対応した４つのドライブ群（第０群：PDEV0～PDEV31、第１群：PDEV32～PDEV63、第２群：PDEV64～PDEV95、第３群：PDEV96～PDEV127）に分割され、各ドライブ群がそれぞれと対応するディスクチャネル４０Ａ～４３Ａ、４０Ｂ～４３Ｂに接続されている。ディスクコントローラ１Ｂも、上記ディスクコントローラ１Ａと同様の構成となっており、ディスクコントローラ１Ｂのディスクアダプタ２０は、ディスクチャネル４Ｂ（４０Ｂ～４３Ｂ）を介してディスクアレイ５に接続されている。

【００３２】

キャッシュメモリ３０は、ディスクコントローラ１Ａ、１Ｂ内の相互結合網４０と接続され、ディスクコントローラ１Ａ、１Ｂの双方からアクセス可能となっている。

【００３３】

図２は、チャネルアダプタ１０の構成を示す。

チャネルアダプタ１０は、チャネルパス３Ａに接続されたホストチャネル・インタフェース１１と、相互結合網４０に接続されたキャッシュメモリ・インタフェース１２と、ＳＶＰ２に接続するためのネットワークインタフェース１３と、ＣＰＵとの間でのデータ転送を制御するためのプロセッサ１４と、該プロセッサが参照する各種のテーブルや実行すべきソフトウェアを格納したローカルメモリ１５と、これらの要素間の相互接続するプロセッサ周辺制御部１６からなる。

【００３４】

ホストチャネル・インタフェース１１は、チャネルパス３Ａ上の制御プロトコルとディスクコントローラ内部の制御プロトコルとの間の変換機能を有し、ホストチャネル・インタフェース１１とキャッシュメモリ・インタフェース１２との間は信号線１７によって接続されている。

【００３５】

図３は、ディスクアダプタ２０の構成を示す。

ディスクアダプタ２０は、相互結合網４０に接続されたキャッシュメモリ・インタフェース２１と、ディスクチャネル４Ａに接続されたディスクチャネル・インタフェース２２と、ＳＶＰ２に接続するためのネットワークインタフェース２

３と、プロセッサ２４と、該プロセッサが参照する各種のテーブルや実行すべきソフトウェアを格納したローカルメモリ２５と、これらの要素間の相互接続するプロセッサ周辺制御部２６とからなる。

【００３６】

キャッシュメモリ・インタフェース２１とディスクチャネル・インタフェース２２との間は信号線２７によって接続されている。ディスクチャネル・インタフェース２２は、ディスクコントローラ内部の制御プロトコルと、ディスクチャネル４Ａ上の制御プロトコル、例えば、ＳＣＳＩとの間の変換機能を備えている。

【００３７】

図４は、本発明によるディスクアレイ５の実装構造の1例を示す。

ディスクアレイ５は、マザーボード（共通基板）６上のコネクタ群８に着脱自在に搭載された複数のドライブ基板BRD（BRD0～BRD31）からなり、本実施例では、各ドライブ基板BRDに４つのディスクドライブPDEV(i)～PDEV(i+3)が搭載されている。上記マザーボード６には、実際には、上記ドライブ基板の他に、例えば、チャネルアダプタ１０やディスクアダプタ２０を構成する各種のＬＳＩチップが搭載された制御基板や、キャッシュメモリ３０用のメモリＬＳＩを搭載したメモリ基板など、他の複数の基板が搭載されるが、ここでは、ドライブ基板のみに着目して説明する。

【００３８】

マザーボードの左上を基点としてディスクアレイ５にＸ、Ｙ、Ｚ軸を設定し、Ｘ軸に沿って配列されたドライブ基板に着目すると、最上段（Ｙ＝０）のドライブ基板BRD0～BRD7に搭載されたディスクドライブPDEV0～PDEV31は、図１に示した第０群のディスクドライブに相当している。これらの第０群のディスクドライブは、ポートバイパス回路（ＰＢＣ）７－０Ａ～７－７Ａ、７－０Ｂ～７－７Ｂを介して、マザーボード６上の第０ディスクチャネル４０Ａ、４０Ｂに接続されている。

【００３９】

同様に、第２段（Ｙ＝１）のドライブ基板BRD8～BRD15に搭載された第１群のディスクドライブPDEV32～PDEV63は、マザーボード６上の第１ディスクチャネル

４１Ａ、４１Ｂに接続され、第３段（Ｙ＝２）のドライブ基板BRD16～BRD23に搭載された第２群のディスクドライブPDEV64～PDEV95は、マザーボード６上の第２ディスクチャネル４２Ａ、４２Ｂに、第４段（Ｙ＝３）のドライブ基板BRD24～BRD31に搭載された第３群のディスクドライブPDEV96～PDEV127は、マザーボード６上の第３ディスクチャネル４３Ａ、４３Ｂに接続されている。

【００４０】

上記ディスクアレイ５において、左端でＹ軸に沿って配列された４枚のドライブ基板BRD0、BRD8、BRD16、BRD24に破線で示したように、互いに同一のＸ座標値、Ｚ座標値をもつ縦方向に配列された４つのディスクドライブによって、ＲＡＩＤグループが形成される。本実施例では、ドライブ基板BRD0、BRD8、BRD16、BRD24上のディスクドライブで形成されるＲＡＩＤグループをＺ座標順にVDEV0～VDEV3と呼ぶ。以下、Ｘ座標の方向にＺ座標順に、ＲＡＩＤグループをVDEV4、VDEV5、VDEV6、VDEV7、･･･VDEV31と呼び、これらのＲＡＩＤグループ名に付された数字０、１、２、･･･をＲＡＩＤグループ識別子と定義する。

【００４１】

各ＲＡＩＤグループ内では、例えば、各ディスクドライブのメモリ空間を複数のストライピング領域に分割し、ストライピング領域によって誤り訂正情報（以下、パリティで代表させる）格納用のディスクドライブが異なるように、パリティ用ドライブの分散化が図られている。

【００４２】

実際の応用においては、ディスクアレイ５に含まれる全てのディスクドライブがＲＡＩＤグループに組み込まれる訳ではなく、そのうちの何個かは予備ドライブとして待機状態となる。以下の説明では、ディスクアレイ５の右端に位置した縦２列のドライブ基板BRD6～BRD30とBRD7～BRD31を予備基板とする。現用系の何れかのドライブ基板でディスクドライブに障害が発生した時、上記予備基板のうちの１つが障害基板に代わる交替基板に選択され、交替基板上のディスクドライブが現用系のＲＡＩＤグループに組み込まれる。

【００４３】

図５と図６は、本発明のディスクアレイ５において、現用系ディスクドライブ

のうちの１つ、例えば、ディスク基板BRD0上の左から２番目のディスクドライブに障害が発生した場合のＲＡＩＤグループ再編成方法を説明するための図である。

【００４４】

本発明のディスクアレイ５では、同一ディスク基板上に搭載された複数のディスクドライブが、それぞれ別々のＲＡＩＤグループに所属しているため、何れかのディスクドライブに障害が発生した時、障害ドライブを含むディスク基板をマザーボードから外す前に、同一ディスク基板上に搭載された正常なディスクドライブについても、ＲＡＩＤグループを再編成しておく。

【００４５】

図５において、ディスク基板BRD0上でドライブ番号「１」をもつディスクドライブが障害ドライブPDEV(E)となった時、PDEV(E)が所属するＲＡＩＤグループVDEV1が異常ＲＡＩＤグループとなる。この場合、ディスク基板BRD0上に搭載されたドライブ番号「０」、「２」、「３」をもつ正常ディスクドライブPDEV(EZ1)、PDEV(EZ2)、PDEV(EZ3)は、障害基板BRD0をマザーボード６から抜き取った時、それぞれが所属する正常なＲＡＩＤグループVDEV0、VDEV2、VDEV3から除外されてしまうため、これらの正常ディスクドライブPDEV(EZ1)、PDEV(EZ2)、PDEV(EZ3)についてもＲＡＩＤグループの再編成が必要となる。本明細書では、障害基板の抜き取りに伴ってＲＡＩＤグループ再編成の対象となる正常ディスクドライブを連累ドライブと呼ぶことにする。

【００４６】

本発明の特徴は、障害回復処理において、障害ドライブPDEV(E)のみならず、連累ドライブPDEV(EZ1)、PDEV(EZ2)、PDEV(EZ3)についても、ＲＡＩＤグループの再編成を行うことにある。このため、予備基板群BRD6～BRD30、BRD7～BRD31の中から、例えば、障害基板BRD0と同一のＹ座標値をもつ予備基板BRD7を交替基板として選択し、この交替基板上の各ディスクドライブに、上記障害基板上の各ドライブの格納データと同一のデータを移し替える。

【００４７】

本実施例の場合、障害回復処理の結果、例えば、図６に示すように、ＲＡＩＤ

グループVDEV0～VDEV3が交替基板上のディスクドライブPDEV(RZ0)～PDEV(RZ3)を含む形に再編成される。ここで、障害ドライブPDEV(E)に代わる交替ドライブPDEV(RZ1)に対しては、ＲＡＩＤグループVDEV1を構成している他の正常ドライブ、ここではドライブ番号「３３」、「６５」、「９７」をもつディスクドライブから読み出されたデータに基づいて、障害ドライブPDEV(E)で失ったデータを回復した後、順次にデータ書込みが行われる。一方、連累ドライブに代わる交替ドライブPDEV(RZ0)、PDEV(RZ2)、PDEV(RZ3)に対しては、連累ドライブを正常にアクセスできるため、それぞれ対応する連累ドライブから読み出したデータを順次に複写すればよい。

【００４８】

図７は、本発明のディスクアレイ５におけるディスクドライブの状態遷移を示す。

現用系の同一ドライブ基板上に搭載された４個のディスクドライブPDEV(EZ0)～PDEV(EZ3)は、最初、正常状態ＳＴ０で動作している。何れかのドライブ、例えば、ディスクドライブPDEV(EZ0)に異常（ＥＶＴ１）が発生すると、障害ドライブPDEV(EZ0)の状態コードは異常状態ＳＴ２に遷移する。

【００４９】

障害基板に代わる交替基板が選択され、ＲＡＩＤグループ（VDEV）の再編成（ＥＶＴ２）が完了すると、障害基板上の各ディスクドライブの状態は交換待ち状態ＳＴ３に遷移する。障害基板がマザーボードから除去され、部品交換した後、マザーボードに正常な基板が挿入（ＥＶＴ３）されると、復旧基板上の各ディスクドライブPDEV(EZ0)～PDEV(EZ3)は、予備状態ＳＴ１となる。これらのディスクドライブは、その後に他のディスクドライブに異常が発生し、ＲＡＩＤグループ（VDEV）の再編成（ＥＶＴ４）で現用系ドライブ群に加えられた時、正常状態ＳＴ０に遷移する。

【００５０】

図８は、ディスク基板上のディスクドライブ群をディスクチャネルに接続するためのポートバイパス回路７の構成を示す。

各ディスクチャネルは、複数のディスクドライブをシリアルに接続したアクセ

スループを形成している。ポートバイパス回路７は、ディスクドライブ（またはドライブ基板）の接続ポートとして機能する回路であって、ディスクチャネル（アクセスループ）の入回線４０Ａからの信号をディスクドライブの入力端ＩＮに供給するためのドライバ回路７１と、上記入回線４０Ａからの入力信号とディスクドライブの出力端ＯＵＴからの信号の何れか一方を選択して、ディスクチャネル（アクセスループ）の出回線４０Ａ’に出力するマルチプレクサ７２からなっている。

マルチプレクサ７２は、例えば、選択信号ＳＥＬＥＣＴが“１”状態の時は、ディスクドライブの出力信号を選択し、選択信号が“０”状態の時は、ディスクチャネルの入回線４０Ａからの入力信号を選択し、出回線４０Ａ’を介して次のポートに供給する。

【００５１】

図９は、第０群のディスクドライブPDEV0～PDEV31と、第０ディスクチャネル４０Ａ、４０Ｂとの接続関係と、バイパス制御信号線SEL-0A～SEL-7A、SEL-0B～SEL-7Bを示す。

マザーボード６上に配線された第０ディスクチャネル４０Ａ、４０Ｂは、第０群のドライブ基板BRD0～BRD7との接続コネクタの近傍に位置して、それぞれポートバイパス回路７－０Ａ～７－７Ａ、７－０Ｂ～７－７Ｂを備えており、これらのポートバイパス回路を介して、ドライブ基板BRD0～BRD7上のアクセスループ４０Ａ－０～４０Ａ－７、４０Ｂ－０～４０Ｂ－７に接続されている。これらのポートバイパス回路は、選択信号ＳＥＬ（SEL-0A～SEL-7A、SEL-0B～SEL-7B）によって制御されている。

【００５２】

一方、ドライブ基板BRD0は、マザーボード上のポートバイパス回路７－０Ａの入出力端子ＩＮとＯＵＴとの間に接続されたアクセスループ４０Ａ－０と、ポートバイパス回路７－０Ｂの入出力端子ＩＮとＯＵＴとの間に接続されたアクセスループ４０Ｂ－０を有し、ドライブ基板BRD0に搭載されたディスクドライブPDEV0～PDEV3は、それぞれ１対のポートバイパス回路（７０－０Ａ、７０－０Ｂ）～（７０－３Ａ、７０－３Ｂ）を備え、ポートバイパス回路７０－０Ａ～７０－３

Ａはアクセスループ４０Ａ－０に接続され、ポートバイパス回路７０－０Ｂ～７０－３Ｂはアクセスループ４０Ｂ－０に接続されている。

【００５３】

ドライブ基板BRD0上のポートバイパス回路（７０－０Ａ、７０－０Ｂ）～（７０－３Ａ、７０－３Ｂ）のＳＥＬＥＣＴ線は、それぞれドライブPDEV0～PDEV3内のポートバイパス制御線と接続され、何れかのドライブが休止状態にある時、該ドライブがアクセスループ４０Ａ－０、４０Ｂ－０からバイパスされるようになっている。

他のドライブ基板BRD1～BRD7上のディスクドライブも、上記ドライブ基板BRD0と同様の構成で、基板上の１対のアクセスループ４０Ａ－ｊ、４０Ｂ－ｊ（ｊ＝１～７）に接続されている。

【００５４】

上記構成から明らかなように、第０群に属したディスクドライブPDEV0～PDEV31は、ドライブ基板およびマザーボードに形成されたディスクチャネル回線（アクセスループ）４０Ａ、４０Ｂによって、互いに直列に接続されており、ディスクドライブに障害が発生した時、障害基板との接続ポートとなるポートバイパス回路の選択信号を“０”にすることによって、障害のあるドライブ基板をディスクチャネル回線４０Ａ、４０Ｂから電気的に分離（バイパス）できるようになっている。

【００５５】

第１０図は、本発明のディスクアレイ装置において、ドライブ基板と該ドライブ基板上に搭載されたディスクドライブとの対応関係を示すために参照される基板管理テーブル８０の１例を示す。

基板管理テーブル８０は、ドライブ基板の識別番号（ＢＲＤ）８０１をもつ複数のデータエントリ８００－０～８００－３１からなる。各データエントリは、ドライブ基板識別番号（ＢＲＤ）８０１と、図４に示したディスクアレイ構成における基板位置を示すＸ座標値８０２およびＹ座標値８０３と、基板の状態コード８０４と、搭載されているディスクドライブ（ＰＤＥＶ）の番号８０５との関係を示している。

【００５６】

状態コード８０４の値ｉは、図７で示した搭載ディスクドライブの状態遷移ＳＴｉ（ｉ＝０～３）に応じて書き換えられる。従って、ディスクアレイに障害ドライブが１つも存在しない場合は、図１０に示すように、状態コード８０４は、現用系として正常に動作中（状態ＳＴ０）であることを示す「０」、または予備ボード（状態ＳＴ１）であることを示す「１」の何れかの値となっている。

【００５７】

第１１図は、本発明のディスクアレイ装置において、障害回復処理のために参照されるドライブ管理テーブル８１の１例を示す。

ドライブ管理テーブル８１は、ディスクドライブ識別番号（ＰＤＥＶ）８１１をもつ複数のデータエントリ８１０－０～８１０－１２７からなる。各データエントリは、ディスクドライブ識別番号（ＰＤＥＶ）８１１と、ディスクドライブが搭載されているドライブ基板の番号８１２と、ドライブ基板上でのディスクドライブの搭載位置を示すＺ座標値８１３と、ディスクドライブのメモリ領域の最大論理ブロックアドレス８１４と、ディスクドライブの状態コード８１５との関係を示している。状態コード８１５の値ｉも、図７で示した状態遷移ＳＴｉ（ｉ＝０～３）に応じて書き換えられる。

【００５８】

第１２図は、本発明のディスクアレイ装置において、ＲＡＩＤグループの構成要素を定義したＲＡＩＤグループ管理テーブル８２の１例を示す。

ＲＡＩＤグループ管理テーブル８２は、ＲＡＩＤグループの識別番号（ＶＤＥＶ）８２１をもつ複数のデータエントリ８２０－０、８２０－１、…からなる。各データエントリは、ＲＡＩＤグループ識別番号（ＶＤＥＶ）８２１と、マスタディスクアダプタ識別番号８２２と、ＲＡＩＤタイプ８２３と、ＲＡＩＤグループの構成ドライブを示すコンポーネントＰＤＥＶ８２４と、ＲＡＩＤグループが再編成された時、新たなＲＡＩＤグループの構成ドライブを示す新コンポーネントＰＤＥＶ８２５との関係を示している。

【００５９】

ここで、マスタディスクアダプタ識別番号８２２は、図１に示すように、ディ

スクアレイ５に２つのディスクコントローラ１Ａ、１Ｂを接続したシステム構成において、マスタとなるディスクコントローラ側のディスクアダプタ２０を示している。

【００６０】

マスタディスクアダプタ識別番号８２２が「０」に設定されたＲＡＩＤグループは、ディスクコントローラ１Ａ側のディスクアダプタ２０がマスタとなり、「１」に設定されたＲＡＩＤグループは、ディスクコントローラ１Ｂ側のディスクアダプタ２０がマスタとなる。マスタに指定されたディスクアダプタ２０に障害が発生した場合は、他方のディスクアダプタ２０が、障害ディスクアダプタに代わってディスクアレイ５をアクセスする。

【００６１】

ＲＡＩＤタイプ８２３は、ＲＡＩＤグループに対して適用されるＲＡＩＤのタイプを示す。図示した例では、ＲＡＩＤグループVDEV0～VDEV8にはＲＡＩＤ５が適用され、VDEV14、VDEV15にはＲＡＩＤ１が適用されていることを示している。

【００６２】

以下、本発明のディスクアレイ装置におけるドライブ異常の回復処理について説明する。

図１３は、現用系のディスクドライブに異常が検出された時、ディスクアダプタ２０のプロセッサ２４が実行する障害回復処理ルーチン２００の１実施例を示す。

【００６３】

図１０～図１２に示した管理テーブル８０～８２は、プロセッサ２４がアクセスするローカルメモリ２５内に形成され、障害回復処理ルーチン２００の実行中に参照される。但し、これらのテーブルはキャッシュメモリ３０に格納されてもよい。また、これらの管理テーブル８０～８２の内容は、ＳＶＰ２を介して保守員に監視され、新エントリの設定とエントリ内容の変更が随時に行なわれる。

【００６４】

プロセッサ２４は、例えば、所定回数のリトライを繰り返してもデータが読み出せない等、現用系のディスクドライブに異常が検出された時、ＳＶＰ２に異常

発生を通知した後、障害回復処理ルーチン２００を実行する。

プロセッサ２４は、先ず、READ/WRITEアクセス中に障害が発生した異常ドライブPDEV(E)の識別番号ｊに基づいて、ドライブ管理テーブル８１から上記異常ドライブと対応するデータエントリ８１０－ｊを検索し、該データエントリ８１０－ｊの状態コード８１５の値を異常状態（ＳＴ２）を示す「2」に変更する。また、上記データエントリ８１０－ｊから、異常ドライブが搭載されている障害基板BRD(E)の識別番号８１２の値ｋを特定する（ステップ２１１）。

【００６５】

プロセッサ２４は、上記障害基板BRD(E)の識別番号ｋに従って基板管理テーブル８０を参照し、データエントリ８００－ｋから、障害基板BRD(E)のＸ座標８０２の値Ex、Ｙ座標８０３の値Eyと、ドライブ番号欄８０５の内容から連累ドライブ番号PDEV(EZ1)～PDEV(EZ3)を特定する（ステップ２１２）。次に、基板管理テーブル８０から、Ｙ座標８０３の値がEyで状態コード８０４が予備状態値「１」となっているデータエントリ８００－ｙを検索し、障害基板BRD(E)に代わるべき交替基板BRD(R)の識別番号ｙと、交替ドライブPDEV(RZ0)～PDEV(RZ3)の識別番号を特定する（ステップ２１３）。また、ＲＡＩＤグループ管理テーブル８２から、コンポーネントPDEV８２４に上記異常ドライブPDEV(E)の識別番号を含むデータエントリと、ステップ２１２で特定された連累ドライブ番号PDEV(EZ1)～PDEV(EZ3)の識別番号を含むデータエントリを検索する。これによって、再編成が必要となるグループ識別子VDEV(N1)～VDEV(N4)が特定される（ステップ２１４）。

【００６６】

以下の説明では、再編成対象となったＲＡＩＤグループのＲＡＩＤタイプ８２３が全て「５」（レベル５）であったと仮定する。

プロセッサ２４は、ＲＡＩＤグループ管理テーブル８２において、ステップ２１４で特定したグループ識別子VDEV(N1)～VDEV(N4)をもつデータエントリの新コンポーネントPDEV８２５として、上記異常ドライブPDEV(E)と連累ドライブ番号PDEV(EZ1)～PDEV(EZ3)の識別番号の代わりに、ステップ２１３で特定した交替ドライブPDEV(RZ0)～PDEV(RZ3)の識別番号を含む新たな組み合せ定義する（ステップ２１５）。

【００６７】

この後、プロセッサ２４は、図１４で詳述する交替ドライブPDEV(RZ0)～PDEV(RZ3)のデータ再生／複写処理ルーチン２２０を実行し、障害基板BRD(E)で保持されていた全てのデータを交替基板上のディスクドライブに再現できた時点で、ＲＡＩＤグループ管理テーブル８２の新コンポーネントPDEV８２５の内容をコンポーネントPDEV８２４に書き写す（ステップ２３１）。上記コンポーネントPDEV８２４の内容変更によって新コンポーネントPDEVのデータが有効化され、図７で示したＲＡＩＤグループPDEVの再編成ＥＶＮＴ２が完了する。

【００６８】

プロセッサ２４は、PDEVの再編成の後処理として、ドライブ管理テーブル８１において、交替ドライブPDEV(RZ0)～PDEV(RZ3)に対応するデータエントリの状態コード８１５の値を正常状態を示す「０」に変更し、異常ドライブPDEV(E)と連累ドライブ番号PDEV(EZ1)～PDEV(EZ3)に対応するデータエントリの状態コード８１５の値を交換待ち状態ＳＴ３を示す「３」に変更する（ステップ２３２）。また、基板管理テーブル８０において、交替基板BRD(R)に対応するデータエントリの状態コード８０４の値を正常状態を示す「０」に変更し、障害基板BRD(E)に対応するデータエントリの状態コード８０４の値を交換待ち状態を示す「３」に変更する（ステップ２３３）。

【００６９】

プロセッサ２４は、上述したテーブル内容の更新が完了すると、障害基板BRD(E)の座標値Ex、Eyで特定される基板接続用のポートバイパス回路の選択信号を切替え、障害基板をディスクチャネル（アクセスループ）から切離し（ステップ２３４）、ＳＶＰ２に対して、障害ドライブと障害基板を特定して、障害回復処理の完了と障害部品の交換を要求する制御メッセージを通知する（ステップ２３５）。

【００７０】

上記制御メッセージの内容は、ＳＶＰ２の表示装置９に出力される。従って、保守員は、上記制御メッセージに従って、障害ボードをマザーボードから取り外し、障害ドライブを正常ドライブに取り替え、部品交換後のドライブ基板をマザ

ーボードに再接続する。この部品交換作業中に、ディスクアレイは再編成されたＲＡＩＤグループによって正常に動作しているため、保守員が部品交換を特に急いで行う必要はない。

【００７１】

保守員は、正常基板をマザーボードに再接続した後、表示装置９に付随する入力装置を利用して、ディスクアダプタ２０のプロセッサ２４に、障害基板の復旧を指令する。プロセッサ２４は、上記指令に応答して、基板管理テーブル８０の復旧基板と対応するエントリの状態コード８０４と、ドライブ管理テーブル８１における上記復旧基板上の搭載ドライブと対応するエントリの状態コード８１５をそれぞれ交換待ち状態値「３」から予備状態値「１」に変更する。また、上記復旧基板BRD(E)の座標値Ex、Eyで特定されるポートバイパス回路の選択信号の状態を切替え、復旧基板上の搭載ドライブをディスクチャネルに接続する。

【００７２】

図１４は、データ再生／複写処理ルーチン２２０の詳細フローチャートの１例を示す。

データ再生／複写処理ルーチン２２０では、先ず、異常ドライブPDEV(E)に格納されていたデータを交替基板BRD(R)上で上記異常ドライブと同一のＺ座標値sをもつ交替ドライブPDEV(RZs)に再生するデータ再生処理３００を実行する。データ再生処理３００の詳細については、後で図１６を参照して説明する。

【００７３】

次に、連累ドライブの残り個数をカウントするためのパラメータｉの値をクリア（ステップ２２１）した後、パラメータｉの値をインクリメントする（ステップ２２２）。この後、障害基板BRD(E)上に搭載されている異常ドライブPDEV(E)以外のドライブPDEV(EZ1)～PDEV(EZ3)の中から選択された第ｉ連累ドライブPDEV(EZi)の格納データをブロック単位で次々と読み出し、読み出されたデータを交替基板BRD(R)上で上記第ｉ連累ドライブと同一のＺ座標値ｉをもつ交替ドライブPDEV(RZi)に複写するデータ複写処理ルーチン４００を実行する。データ複写処理４００の詳細については、後で図１７を参照して説明する。

【００７４】

データ複写処理ルーチン４００が完了すると、パラメータｉの値が連累ドライブの基板搭載個数（この例では３個）に達したか否かを判定し（ステップ２２３）、連累ドライブの基板搭載個数に達していれば、このルーチン２２０を終了し、そうでなければ、ステップ２２２に戻って、次の連累ドライブについて同様の処理を繰り返す。

【００７５】

図１５は、データ再生処理３００において、プロセッサ２４が利用するデータ再生管理テーブル８３の構成を示す。

データ再生処理３００では、図１６で詳述するように、交替ドライブPDEV(RZs)への再生データの書込みが、複数データブロックからなる領域単位で繰り返される。

【００７６】

データ再生管理テーブル８３は、交替ドライブPDEV(RZs)のメモリ空間を複数のメモリ領域に分割して各メモリ領域に割当てた領域番号８３１と、メモリ領域の開始アドレスおよび最終アドレスを示す論理ブロックアドレス８３２と、再生済みフラグ８３３とを示す複数のエントリ８３０－１、８３０－１、…からなっている。再生データ書込み済みとなった領域には、再生済みフラグ８３３に値「１」が設定される。

【００７７】

図１６は、データ再生処理３００の詳細フローチャートの１例を示す。

データ再生処理３００では、プロセッサ２４は、初期値として、パラメータSADDに交替ドライブPDEV(RZs)のメモリ空間における再生開始アドレスを設定し（ステップ３０１）、パラメータEADDに交替ドライブPDEV(RZs)のメモリ空間における再生終了アドレスを設定し（ステップ３０２）、パラメータBLに再生単位となる領域サイズ（またはデータブロック数）を設定し（ステップ３０３）、データ再生領域を切替えるためのパラメータCNTにパラメータSADDの値を設定する（ステップ３０４）。パラメータBLが示す領域サイズは、データ生成管理テーブル８３における各論理ブロックアドレス８３が示すアドレス範囲に相当している。

【００７８】

次に、プロセッサ２４は、パラメータCNTとEADDの値を比較する（ステップ３０５）。CNTの値がEADD以上の場合は、このルーチンを終了する。CNT＜EADDの場合は、異常ドライブPDEV(Eのメモリ空間に対して上位装置からアクセス要求（READ/WRITE命令）があったか否かを判定する（ステップ３０６）。もし、アクセス要求がなければ、データ生成管理テーブル８３を参照し、パラメータCNTで特定されるメモリ領域の再生済みフラグ８３３をチェックする（ステップ３０７）。

【００７９】

再生済みフラグ８３３が「１」となっていた場合、プロセッサ２４は、パラメータCNTにブロックサイズBLの値を加算し（ステップ３１０）、ステップ３０５に戻る。再生済みフラグ８３３が「０」となっていた場合は、既にＲＡＩＤグループ管理テーブル８２のコンポーネントPDEV８２４で特定済みとなっている上記異常ドライブPDEV(E)が所属するＲＡＩＤグループの他の正常ディスクドライブから、パラメータCNTとBLで特定される１メモリ領域分のデータを読み出し、読み出されたデータのビット毎の排他論理和（ＥｘＯＲ）演算結果を交替ドライブPDEV(RZs)の該当メモリ領域に次々と書き込む（ステップ３０８）。

この後、データ再生管理テーブル８３のパラメータCNTで特定されるメモリ領域、すなわち、論理ブロックアドレス８３２がCNT～CNT+BLのデータエントリの再生済みフラグ８３３に「１」を設定し（ステップ３０９）、ステップ３１０を実行する。

【００８０】

判定ステップ３０７で上位装置からアクセス要求を受けていた場合は、データ再生管理テーブル８３を参照し、アクセス要求が示すアドレス範囲がデータ再生済みのメモリ領域に該当しているか否かを判定する（ステップ３２０）。アクセスすべきメモリ領域が、交替ドライブにデータ再生済みのメモリ領域に該当していた場合は、交替ドライブPDEV(RZs)を含む新ＲＡＩＤグループに対してREAD/WRITE命令を実行し（ステップ３２３）、その後でステップ３０６を実行する。

【００８１】

アクセスすべきメモリ領域が、未だデータ再生されていないメモリ領域に該当していた場合は、ＲＡＩＤグループの他の正常ディスクドライブから、アクセス

対象領域のデータを読み出し、読み出されたデータのビット毎の排他論理和（ＥｘＯＲ）演算結果を交替ドライブPDEV(RZs)の該当メモリ領域に次々と書き込む（ステップ３２１）。この後、データ再生管理テーブル８３で上記メモリ領域に該当するデータエントリの再生済みフラグ８３３に「１」を設定し（ステップ３２２）、ステップ３２３で上記交替ドライブPDEV(RZs)を含む新ＲＡＩＤグループに対してREAD/WRITE命令を実行する。

【００８２】

図１７は、交替ドライブへのデータ複写処理４００の詳細フローチャートの１例を示す。

データ複写処理４００において、プロセッサ２４は、ステップ４０１～４０４で、データ再生処理３００のステップ３０１～３０４と同様の初期値設定を行った後、パラメータCNTとEADDの値を比較する（ステップ４０５）。CNTの値がEADD以上となっていた場合は、このルーチンを終了し、CNT＜EADDの場合は、現在、処理対象となっている連累ドライブPDEV(EZi)のメモリ空間に対して上位装置からアクセス要求（READ/WRITE命令）があったか否かを判定する（ステップ４０６）。

【００８３】

もし、アクセス要求がなければ、連累ドライブPDEV(EZi)から、パラメータCNTとBLで特定されるメモリ領域から複数データブロック分のデータを読み出し、これを交替ドライブPDEV(RZi)の該当メモリ領域に書き込む（ステップ４０７）。この後、パラメータCNTにブロックサイズBLの値を加算し（ステップ４０８）、ステップ４０５に戻る。

【００８４】

判定ステップ４０６で上位装置からアクセス要求を受けていた場合は、アクセス要求の種類を判定する（ステップ４１０）。アクセス要求がREAD命令の場合は、連累ドライブPDEV(EZi)のメモリ空間において上記READ命令を実行し（ステップ４１１）、命令実行後、ステップ４０６に進む。

【００８５】

アクセス要求がWRITE命令の場合、アクセスすべきメモリ領域が、交替ドライ

ブにデータ複写済みのメモリ領域に該当しているか否かを判定する（ステップ４１２）。アクセスすべきメモリ領域がデータ複写済みのメモリ領域に該当していた場合は、交替ドライブPDEV(RZi)を含む新ＲＡＩＤグループに対してWRITE命令を実行し（ステップ４１３）、命令実行後、ステップ４０６に進む。アクセスすべきメモリ領域が未だデータ複写が終わっていないメモリ領域に該当していた場合は、連累ドライブPDEV(EZi)を含む元のＲＡＩＤグループに対してWRITE命令を実行する（ステップ４１４）。

【００８６】

図１３では、ＲＡＩＤ５を前提に障害回復処理２００の実施例を説明したが、再編成対象となったＲＡＩＤグループのＲＡＩＤタイプ８２３が、例えば、ＲＡＩＤ１(レベル１)の場合は、ＲＡＩＤグループのコンポーネントが障害ドライブと少なくとも１つの予備ドライブであり、データ再生／複写処理２２０では、予備ドライブの格納データを交替ドライブに複写すれば済む。従って、コンポーネント数が少なくなるだけで、図１３のステップ２１４～２３２と同様の手順で、ＲＡＩＤグループの再編成を実現できる。

【００８７】

以上の実施例では、ドライブ基板上の１つのディスクドライブに障害が発生した場合について説明したが、本発明は、同一基板上の複数のディスクドライブに障害が発生した場合でも適用できる。例えば、同一基板上の２個のディスクドライブに障害が発生した場合は、各障害ドライブに対してデータ再生処理３００を実行し、残りの２個の連累ドライブに対してデータ複写処理４００を実行すればよい。

【００８８】

実施例では、データ再生／複写処理２２０をディスクアダプタ２０のプロセッサ２４で実行したが、データ再生／複写処理２２０の一部をプロセッサ２４からの指令によって、他の処理装置に代行させるようにしてもよい。特に、連累ドライブから交替ドライブへのデータの複写は、ＳＣＳＩにおけるEXTENDED COPYコマンドを利用して実行できる。

【００８９】

従って、例えば、図１８に示すように、ディスクコントローラ１Ａ（１Ｂ）とディスクチャネル４Ａ（４Ｂ）との間にスイッチＳＷ－Ａ（ＳＷ－Ｂ）を配置し、上記スイッチを介して代行処理装置ＰＵ－Ａ（ＳＷ－Ｂ）をディスクチャネル４Ａ（４Ｂ）に接続するシステム構成としておき、代行処理装置が、プロセッサ２４からのEXTENDED COPYコマンドに応答して、指定された連累ドライブの格納データを指定された交替ドライブに複写するようにしてもよい。

【００９０】

【発明の効果】

以上の実施例から明らかなように、本発明によれば、ＲＡＩＤグループの再編成が完了した後に障害基板の部品を交換するようにしているため、障害基板の取り外し期間中でも通常のデータ・リード／ライトが可能となる。従って、本発明によれば、ドライブ基板の実装密度を高くし、且つ、障害ドライブの部品交換に十分な作業時間を許容できる。また、部品交換作業中に不在ドライブに対する論理グループによるデータ再生を行う必要がないため、上位装置からのディスクアクセス要求に対して迅速の応答することが可能となる。

【図面の簡単な説明】

【図１】

本発明が適用されるディスクアレイ装置の１実施例を示す図。

【図２】

チャネルアダプタ１０の構成を示すブロック図。

【図３】

ディスクアダプタ２０の構成を示すブロック図。

【図４】

本発明によるディスクアレイ５の実装構造の１例を示す斜視図。

【図５】

ディスクドライブに異常が発生した場合に本発明で再編成すべきＲＡＩＤグループと交替基板との関係を示す図。

【図６】

再編成後のＲＡＩＤグループの構成を示す図。

【図７】

本発明のディスクアレイにおけるディスクドライブの状態遷移を説明するための図。

【図８】

ポートバイパス回路の構成図。

【図９】

ディスクドライブ群とディスクチャネルとの接続関係を示す図。

【図１０】

障害回復処理で参照される基板管理テーブル８０の構成図。

【図１１】

障害回復処理で参照されるドライブ管理テーブル８１の構成図。

【図１２】

障害回復処理で参照されるＲＡＩＤグループ管理テーブル８２の構成図。

【図１３】

本発明による障害回復処理ルーチン２００の１実施例を示すフローチャート。

【図１４】

図１３の障害回復処理ルーチン２００に含まれるデータ再生／複写処理ルーチン２２０の詳細フローチャート。

【図１５】

データ再生／複写処理ルーチン２２０で参照されるデータ再生管理テーブルの構成図。

【図１６】

データ再生／複写処理ルーチン２２０におけるデータ再生処理３００の詳細フローチャート。

【図１７】

データ再生／複写処理ルーチン２２０におけるデータ複写処理４００の詳細フローチャート。

【図１８】

本発明が適用されるディスクアレイ装置の他の実施例を示す図。

【符号の説明】

１：ディスクコントローラ、２：ＳＶＰ、３：チャネルパス、

４：ディスクチャネル、５：ディスクアレイ、６：共通基板（マザーボード）、

７：ポートバイパス回路、８：コネクタ、１０：チャネルアダプタ、

２０：ディスクアダプタ、１１：ホストチャネルインタフェース、

１２、２１：キャッシュメモリインタフェース、

１３、２３：ネットワークインタフェース、１４、２４：プロセッサ、

１５、２５：ローカルメモリ、２２：ディスクチャネルインタフェース、

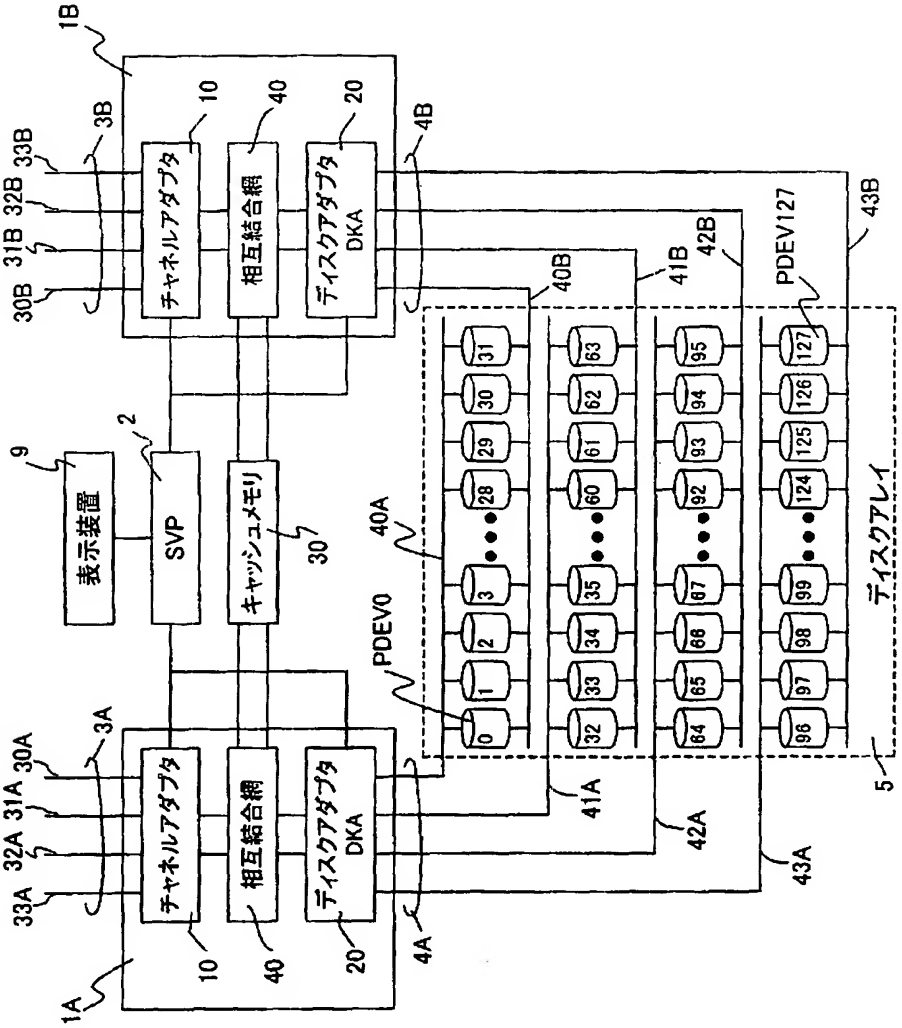
PDEV：ディスクドライブ、VDEV：ＲＡＩＤグループ、BRD：ドライブ基板、

ＳＷ－Ａ、ＳＷ－Ｂ：スイッチ、ＰＵ－Ａ、ＰＵ－Ｂ：代行処理装置。

【書類名】図面

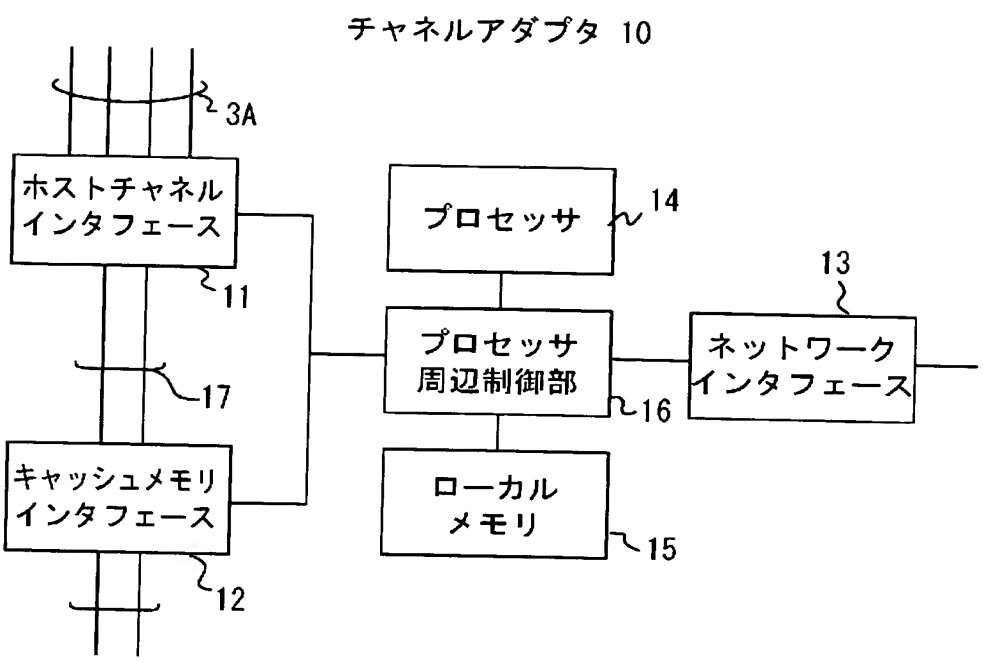
【図１】

図　1


1B
3B
33B
32B
31B
30B
チャネルアダプタ
10
相互結合網
40
ディスクアダプタ
DKA
20
4B
43B
PDEV127
42B
41B
40B
9
表示装置
2
SVP
キャッシュメモリ
30
40A
PDEV0
ディスクアレイ
5
3A
30A
31A
32A
33A
41A
42A
43A
4A
1A

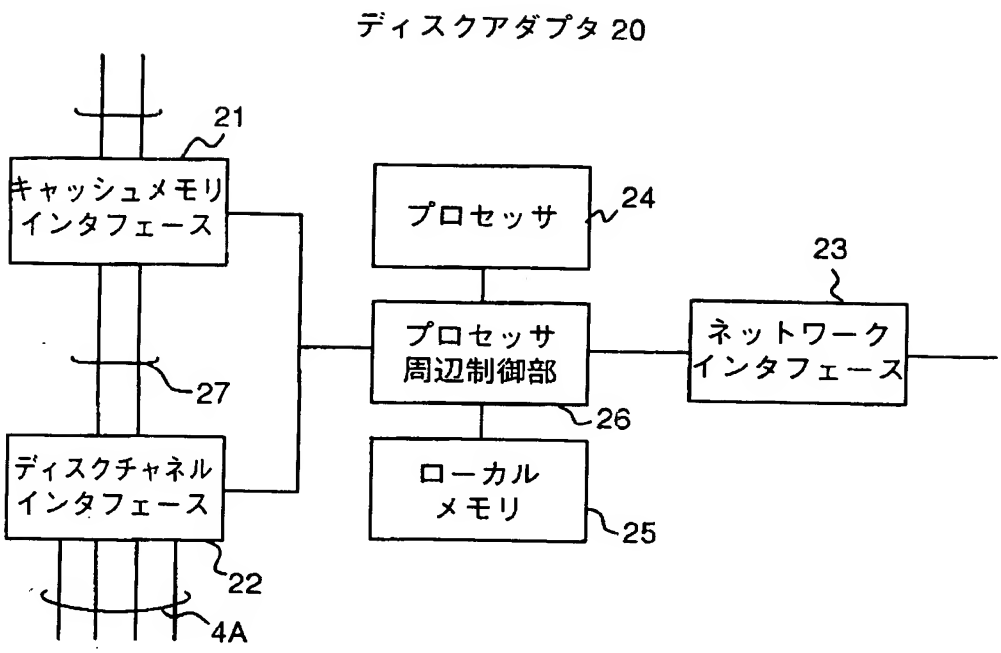
【図２】

図　2


チャネルアダプタ 10
3A
ホストチャネル
インタフェース
11
17
キャッシュメモリ
インタフェース
12
プロセッサ
14
プロセッサ
周辺制御部
16
ローカル
メモリ
15
13
ネットワーク
インタフェース

【図３】

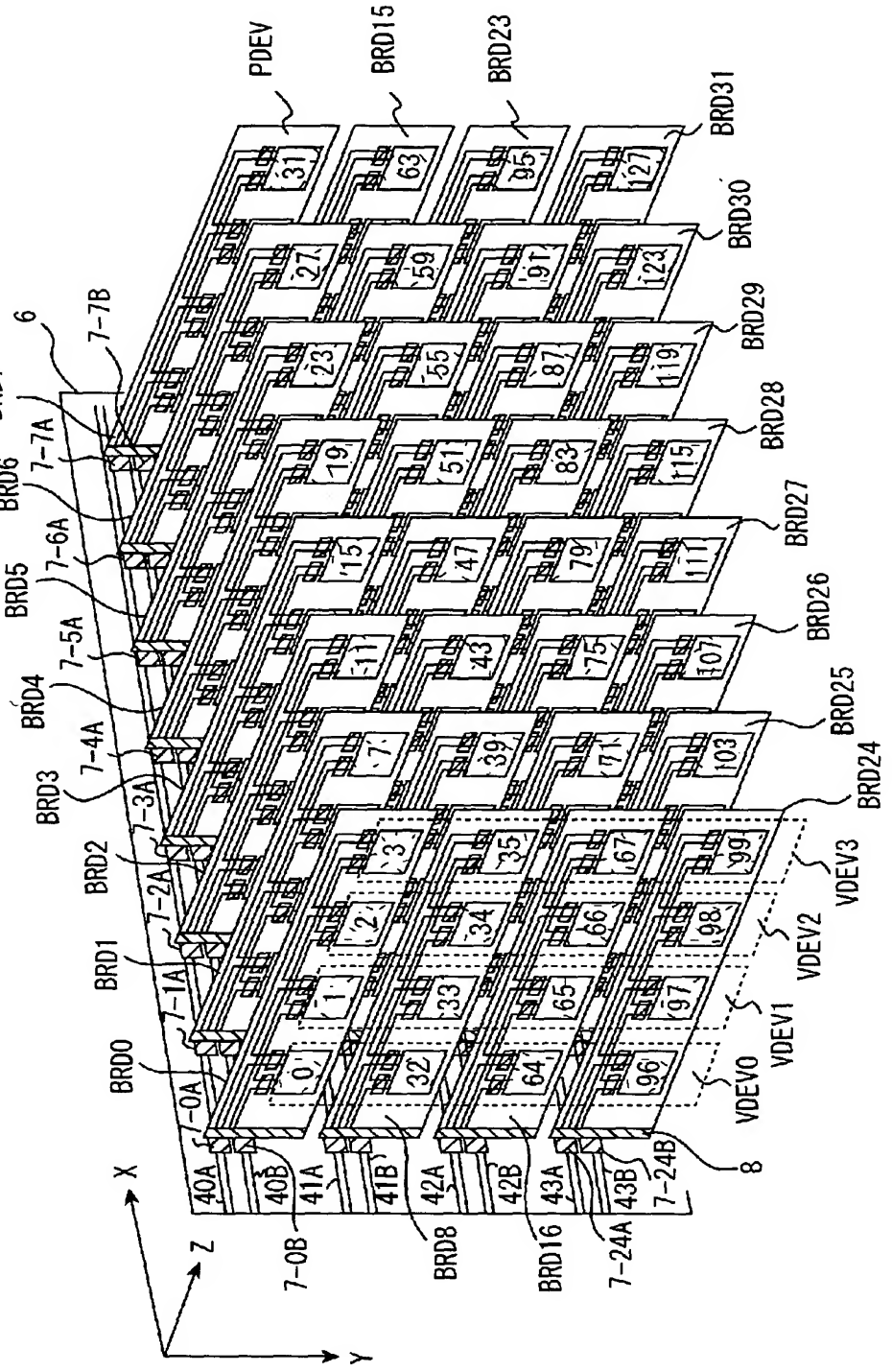
図　3


ディスクアダプタ 20
21
キャッシュメモリ
インタフェース
プロセッサ
24
23
プロセッサ
周辺制御部
ネットワーク
インタフェース
27
26
ディスクチャネル
インタフェース
ローカル
メモリ
25
22
4A

【図４】

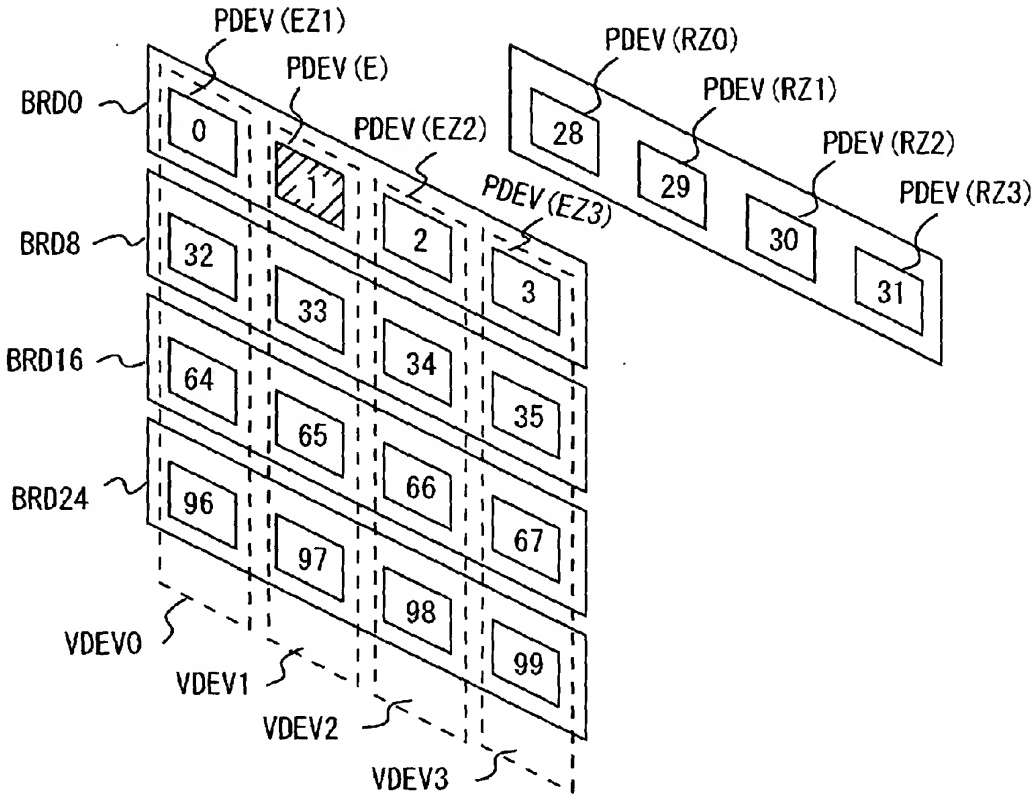
図 4

ディスクアレイ 5



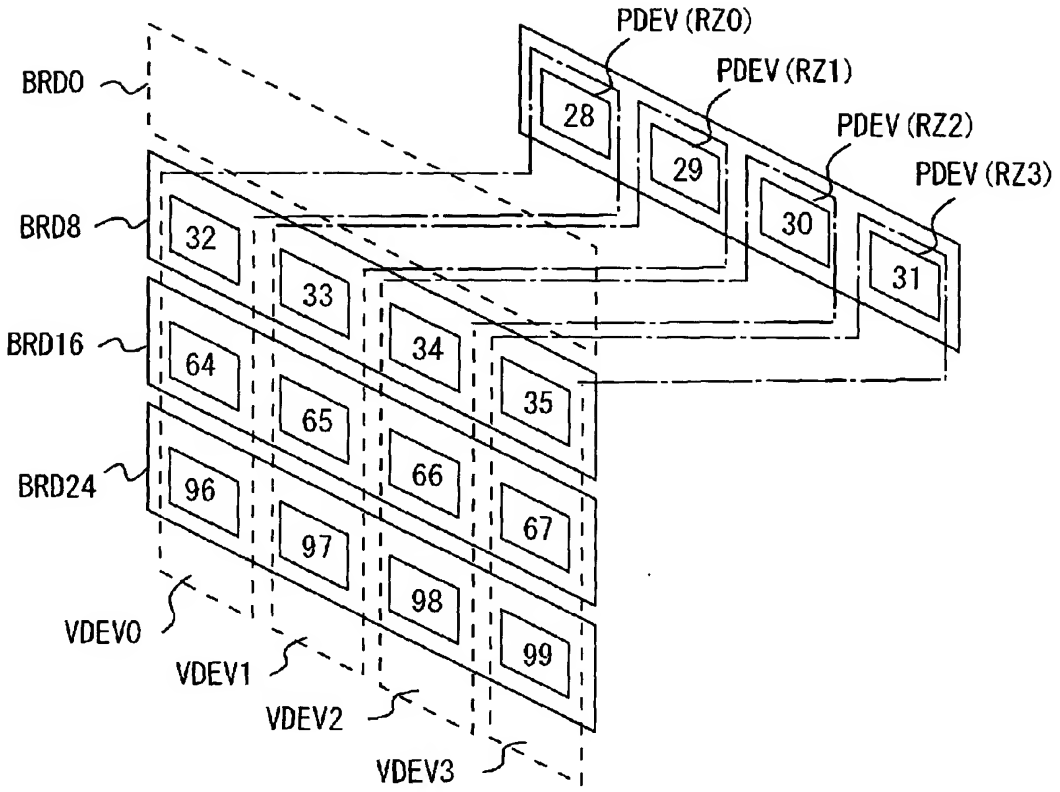
【図５】

図　5


PDEV(EZ1)
PDEV(E)
PDEV(EZ2)
PDEV(EZ3)
PDEV(RZ0)
PDEV(RZ1)
PDEV(RZ2)
PDEV(RZ3)
BRD0
BRD8
BRD16
BRD24
0
1
2
3
28
29
30
31
32
33
34
35
64
65
66
67
96
97
98
99
VDEV0
VDEV1
VDEV2
VDEV3

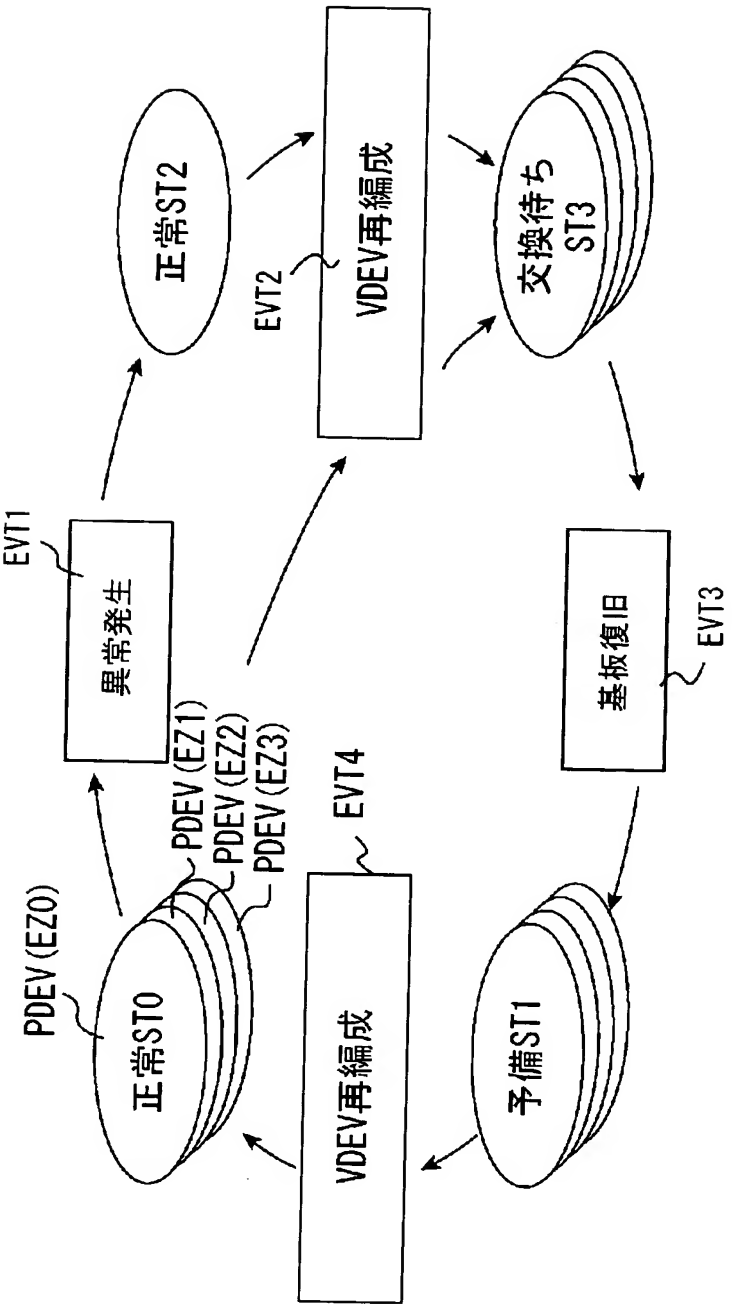
【図６】

図　6


PDEV (RZ0)
PDEV (RZ1)
PDEV (RZ2)
PDEV (RZ3)
BRD0
BRD8
BRD16
BRD24
28
29
30
31
32
33
34
35
64
65
66
67
96
97
98
99
VDEV0
VDEV1
VDEV2
VDEV3

【図７】

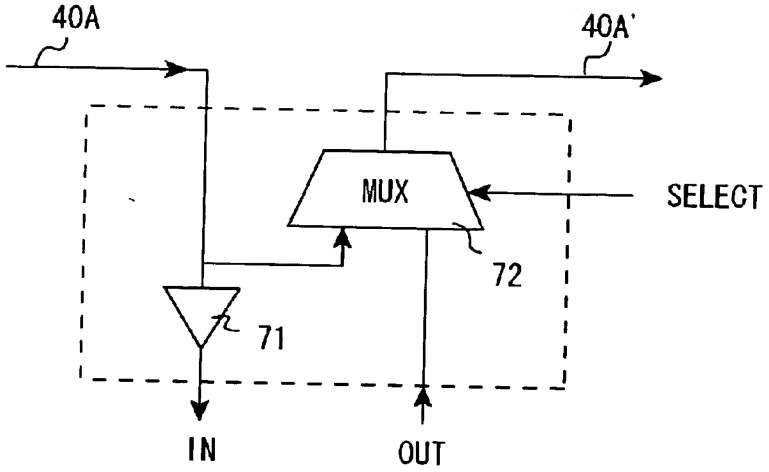
図　7


正常ST2
EVT2
VDEV再編成
交換待ち
ST3
EVT1
異常発生
基板復旧
EVT3
PDEV(EZ1)
PDEV(EZ2)
PDEV(EZ3)
EVT4
PDEV(EZ0)
正常ST0
VDEV再編成
予備ST1

【図８】

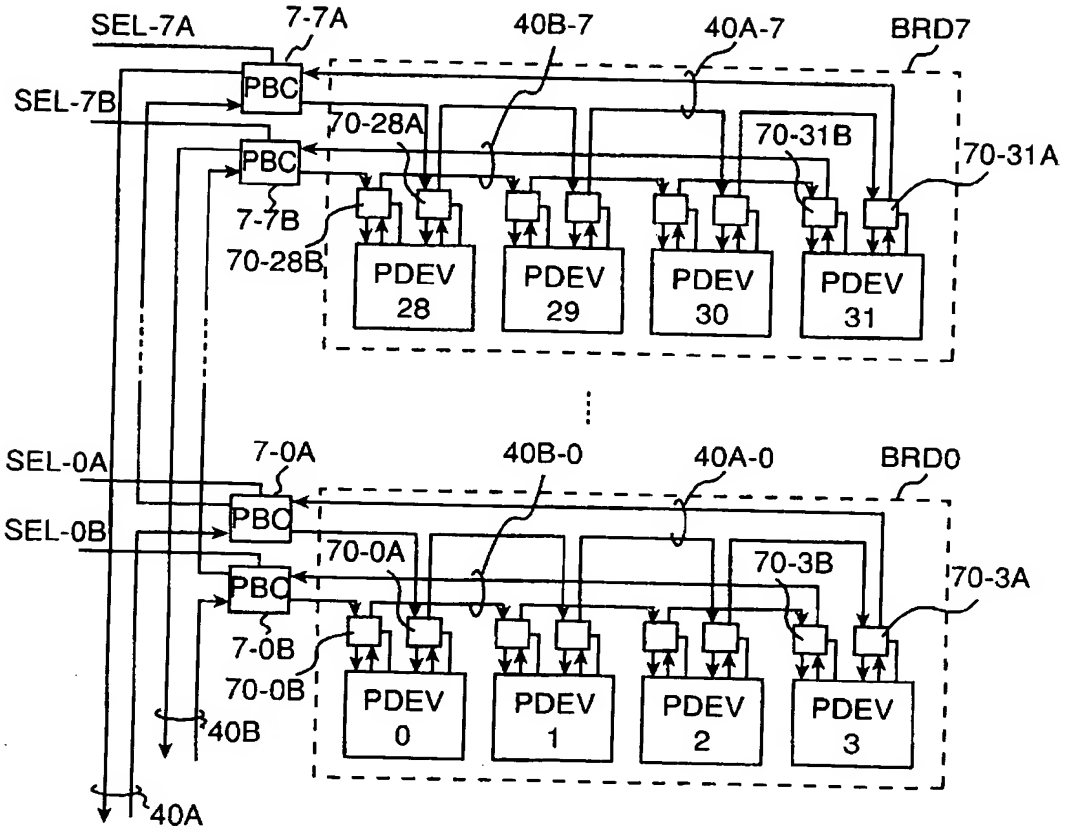
図　8

ポートバイパス回路（PBC）7



【図９】

図　9


SEL-7A
7-7A
40B-7
40A-7
BRD7
SEL-7B
PBC
PBC
70-28A
70-31B
70-31A
7-7B
70-28B
PDEV 28
PDEV 29
PDEV 30
PDEV 31
SEL-0A
7-0A
40B-0
40A-0
BRD0
SEL-0B
PBC
PBC
70-0A
70-3B
70-3A
7-0B
70-0B
PDEV 0
PDEV 1
PDEV 2
PDEV 3
40B
40A

【図１０】

図　10

基板管理テーブル 80

801 基板番号 BRD	802 X座標	803 Y座標	804 基板状態コード	805 ドライブ番号 PDEV	
0	0	0	0	{ 0, 1, 2, 3}	800-0
1	1	0	0	{ 4, 5, 6, 7}	800-1
2	2	0	0	{ 8, 9, 10, 11}	800-2
3	3	0	0	{ 12, 13, 14, 15}	⋮
4	4	0	0	{ 16, 17, 18, 19}	
5	5	0	0	{ 20, 21, 22, 23}	
6	6	0	1	{ 24, 25, 26, 27}	
7	7	0	1	{ 28, 29, 30, 31}	
8	0	1	0	{ 32, 33, 34, 35}	
9	1	1	0	{ 36, 37, 38, 39}	
10	2	1	0	{ 40, 41, 42, 43}	
11	3	1	0	{ 44, 45, 46, 47}	
12	4	1	0	{ 48, 49, 50, 51}	
13	5	1	0	{ 52, 53, 54, 55}	
14	6	1	1	{ 56, 57, 58, 59}	
15	7	1	1	{ 60, 61, 62, 63}	
16	0	2	0	{ 64, 65, 66, 67}	
17	1	2	0	{ 68, 69, 70, 71}	
18	2	2	0	{ 72, 73, 74, 75}	
19	3	2	0	{ 76, 77, 78, 79}	
20	4	2	0	{ 80, 81, 82, 83}	
21	5	2	0	{ 84, 85, 86, 87}	
22	6	2	1	{ 88, 89, 90, 91}	
23	7	2	1	{ 92, 93, 94, 95}	
24	0	3	0	{ 96, 97, 98, 99}	
25	1	3	0	{100, 101, 102, 103}	
26	2	3	0	{104, 105, 106, 107}	
27	3	3	0	{108, 109, 110, 111}	
28	4	3	0	{112, 113, 114, 115}	
29	5	3	0	{116, 117, 118, 119}	
30	6	3	1	{120, 121, 122, 123}	800-31
31	7	3	1	{124, 125, 126, 127}	

【図１１】

図　11

ドライブ管理テーブル 81

811 PDEV	812 基板番号	813 Z座標	814 最大論理ブロックアドレス	815 PDEV状態コード	
0	0	0	0x1FFFFFFF	0	810-0
1	0	1	0x1FFFFFFF	0	810-1
2	0	2	0x1FFFFFFF	0	
3	0	3	0x1FFFFFFF	0	
4	1	0	0x1FFFFFFF	0	
5	1	1	0x1FFFFFFF	0	
6	1	2	0x1FFFFFFF	0	
7	1	3	0x1FFFFFFF	0	
8	2	0	0x1FFFFFFF	0	
9	2	1	0x1FFFFFFF	0	
10	2	2	0x1FFFFFFF	0	
11	2	3	0x1FFFFFFF	0	
12	3	0	0x1FFFFFFF	0	
13	3	1	0x1FFFFFFF	0	
14	3	2	0x1FFFFFFF	0	
15	3	3	0x1FFFFFFF	0	
16	4	0	0x1FFFFFFF	0	
⋮	⋮	⋮	⋮		
32	8	0	0x1FFFFFFF	0	
33	8	1	0x1FFFFFFF	0	
34	8	2	0x1FFFFFFF	0	
35	8	3	0x1FFFFFFF	0	
36	9	1	0x1FFFFFFF	0	
⋮	⋮	⋮	⋮		
122	30	2	0x1FFFFFFF	1	
123	30	3	0x1FFFFFFF	1	
124	31	0	0x1FFFFFFF	1	
125	31	1	0x1FFFFFFF	1	
126	31	2	0x1FFFFFFF	1	
127	31	3	0x1FFFFFFF	1	810-127

【図１２】

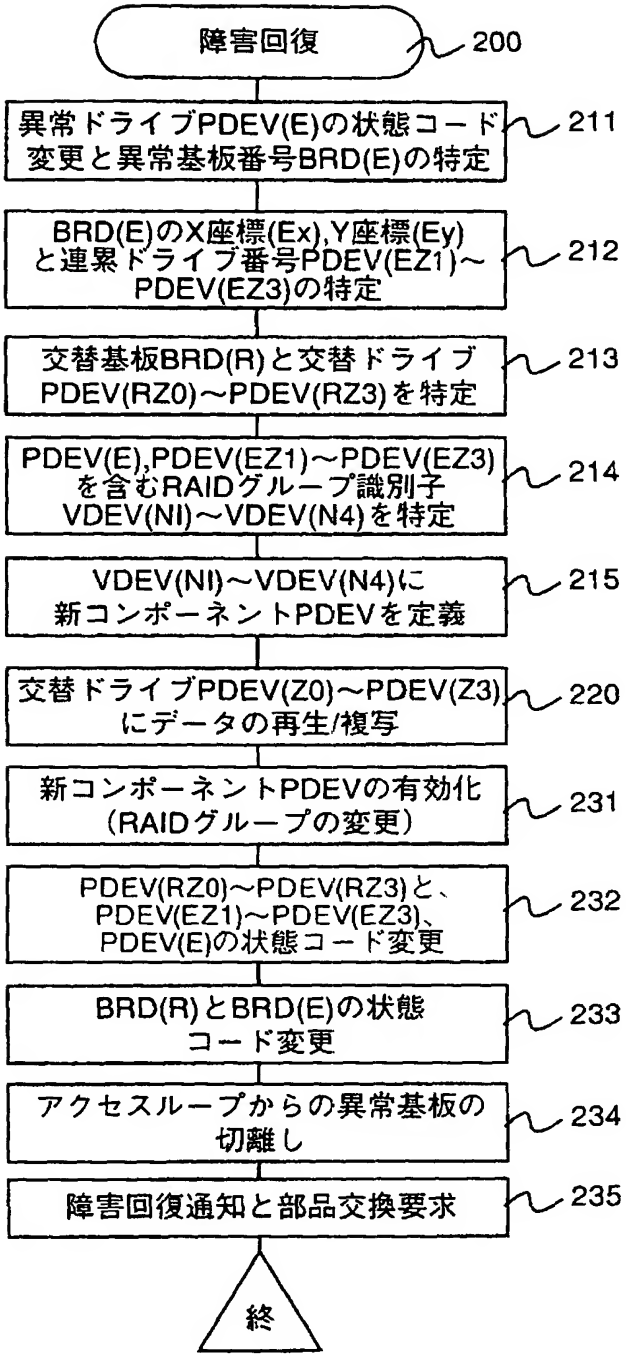
図　12

RAIDグループ管理テーブル 82

821 グループ識別子 VDEV	822 マスタディスクアダプタ番号	823 RAIDタイプ	824 コンポーネント PDEV	825 新コンポーネント PDEV	
0	0	5	{0, 32, 64, 96}	{28, 32, 64, 96}	～820-1
1	0	5	{1, 33, 65, 97}	{29, 33, 65, 97}	～820-2
2	0	5	{2, 34, 66, 98}	{30, 34, 66, 98}	⋮
3	0	5	{3, 35, 67, 99}	{31, 35, 67, 99}	
4	1	5	{4, 36, 68, 100}		
5	1	5	{5, 37, 69, 101}		
6	1	5	{6, 38, 70, 102}		
7	1	5	{7, 39, 71, 103}		
8	0	5	{8, 40, 72, 104}		
⋮	⋮	⋮	⋮		
14	0	1	{14, 46}		
15	0	1	{15, 47}		
⋮	⋮	⋮	⋮		

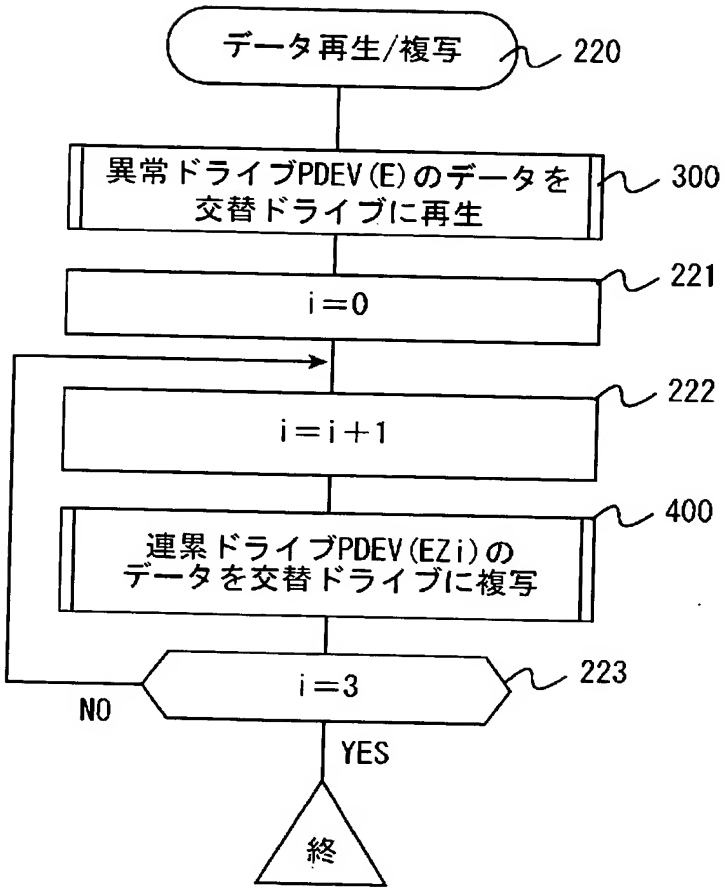
【図１３】

図　13


障害回復
200
異常ドライブPDEV(E)の状態コード変更と異常基板番号BRD(E)の特定
211
BRD(E)のX座標(Ex),Y座標(Ey)と連累ドライブ番号PDEV(EZ1)～PDEV(EZ3)の特定
212
交替基板BRD(R)と交替ドライブPDEV(RZ0)～PDEV(RZ3)を特定
213
PDEV(E),PDEV(EZ1)～PDEV(EZ3)を含むRAIDグループ識別子VDEV(NI)～VDEV(N4)を特定
214
VDEV(NI)～VDEV(N4)に新コンポーネントPDEVを定義
215
交替ドライブPDEV(Z0)～PDEV(Z3)にデータの再生/複写
220
新コンポーネントPDEVの有効化(RAIDグループの変更)
231
PDEV(RZ0)～PDEV(RZ3)と、PDEV(EZ1)～PDEV(EZ3)、PDEV(E)の状態コード変更
232
BRD(R)とBRD(E)の状態コード変更
233
アクセスループからの異常基板の切離し
234
障害回復通知と部品交換要求
235
終

【図１４】

図　14


データ再生/複写
220
異常ドライブPDEV(E)のデータを
交替ドライブに再生
300
i=0
221
i=i+1
222
連累ドライブPDEV(EZi)の
データを交替ドライブに複写
400
i=3
223
NO
YES
終

【図１５】

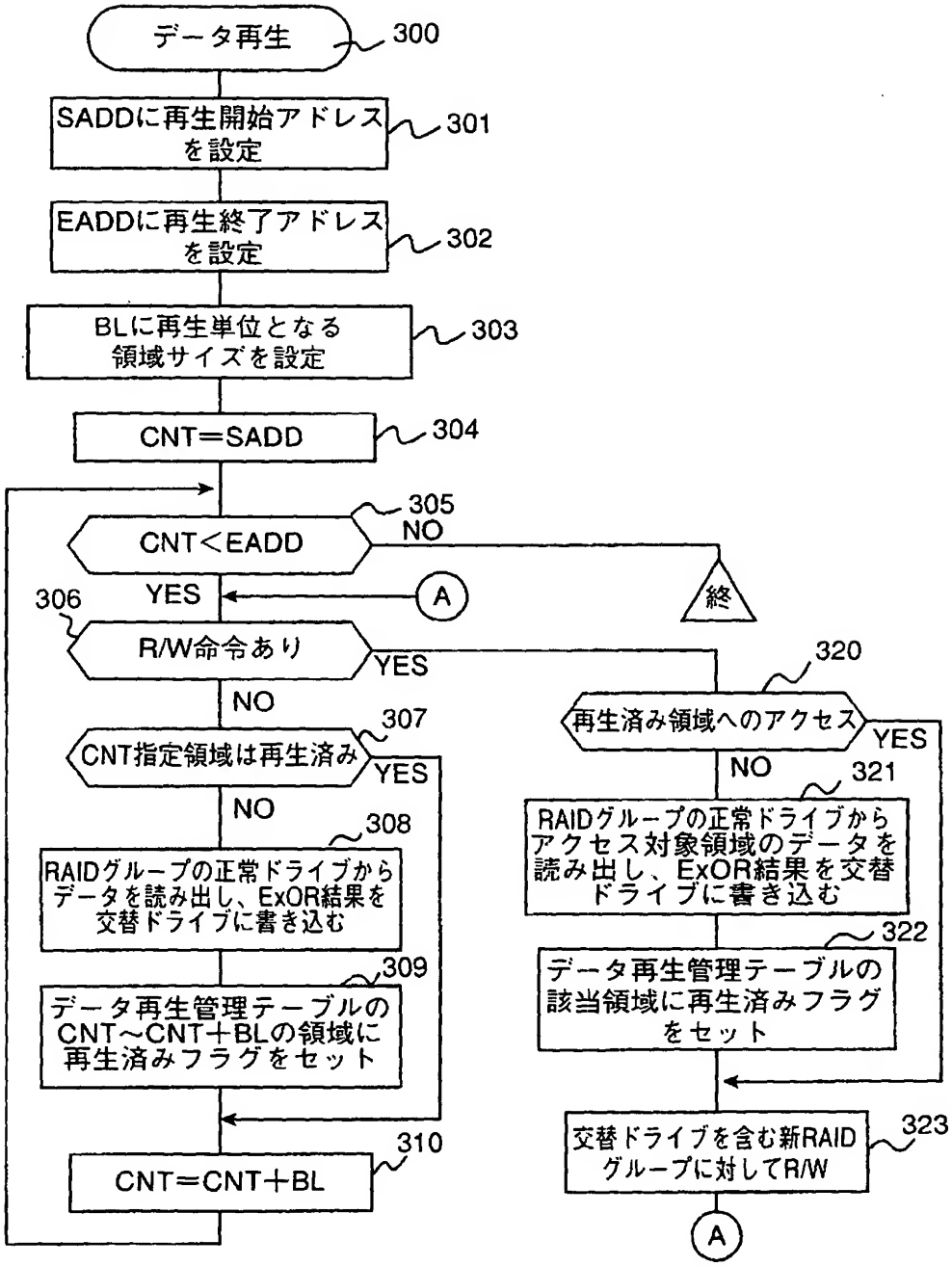
図　15

データ再生管理テーブル 83

831 領域番号	832 論理ブロックアドレス		833 再生済みフラグ	
	From	To		
0	0x00000000	0x0000FFFF	1	830-0
1	0x00010000	0x0001FFFF	1	830-1
2	0x00020000	0x0002FFFF	1	830-2
3	0x00030000	0x0003FFFF	1	
4	0x00040000	0x0004FFFF	0	
5	0x00050000	0x0005FFFF	0	
6	0x00060000	0x0006FFFF	0	
7	0x00070000	0x0007FFFF	0	
8	0x00080000	0x0008FFFF	0	
9	0x00090000	0x0009FFFF	1	
10	0x000A0000	0x000AFFFF	0	

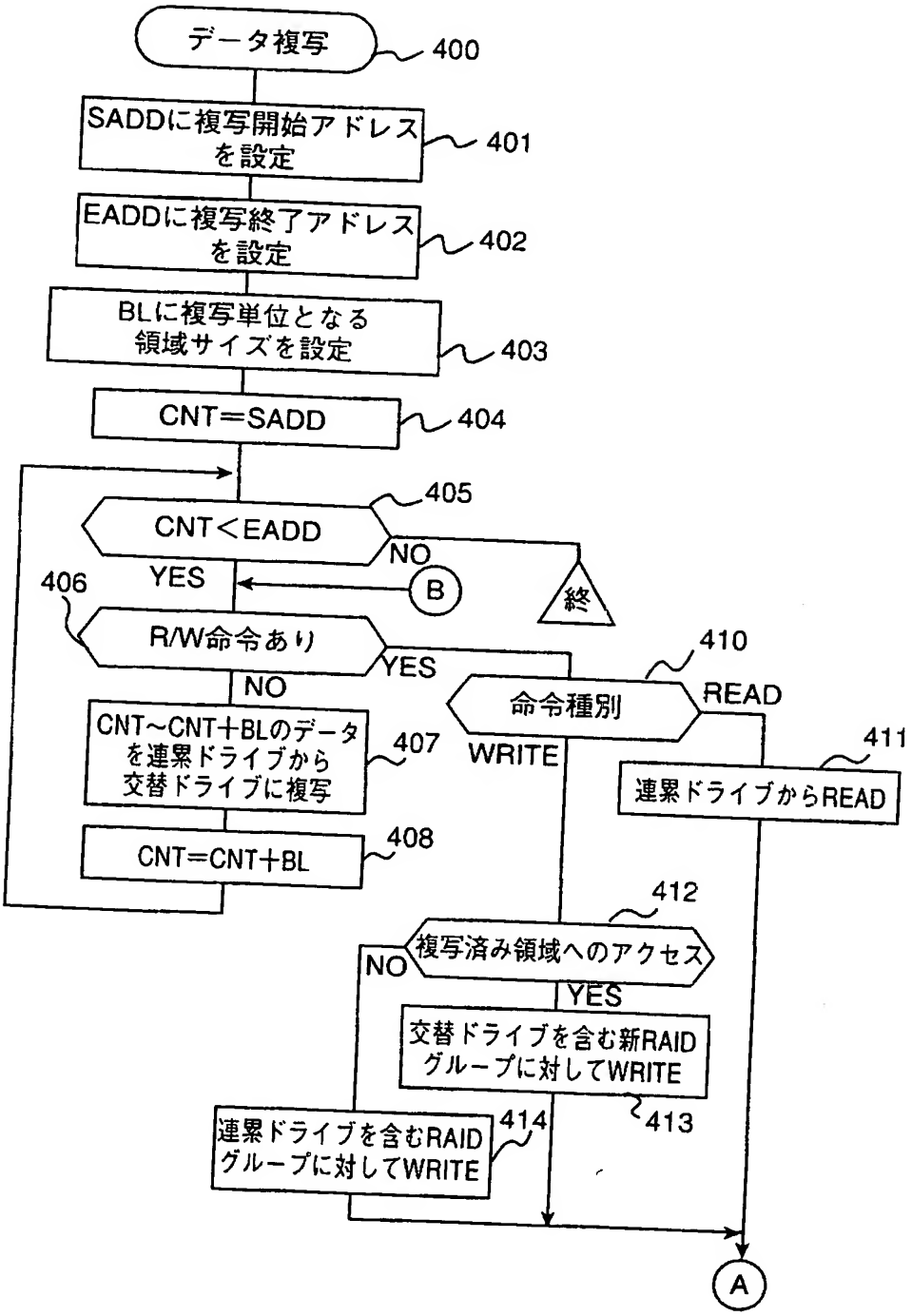
【図１６】

図　16


データ再生
300
SADDに再生開始アドレスを設定
301
EADDに再生終了アドレスを設定
302
BLに再生単位となる領域サイズを設定
303
CNT=SADD
304
305
CNT<EADD
NO
終
YES
A
306
R/W命令あり
YES
320
再生済み領域へのアクセス
YES
NO
321
RAIDグループの正常ドライブからアクセス対象領域のデータを読み出し、ExOR結果を交替ドライブに書き込む
322
データ再生管理テーブルの該当領域に再生済みフラグをセット
NO
307
CNT指定領域は再生済み
YES
NO
308
RAIDグループの正常ドライブからデータを読み出し、ExOR結果を交替ドライブに書き込む
309
データ再生管理テーブルのCNT～CNT+BLの領域に再生済みフラグをセット
310
CNT=CNT+BL
323
交替ドライブを含む新RAIDグループに対してR/W
A

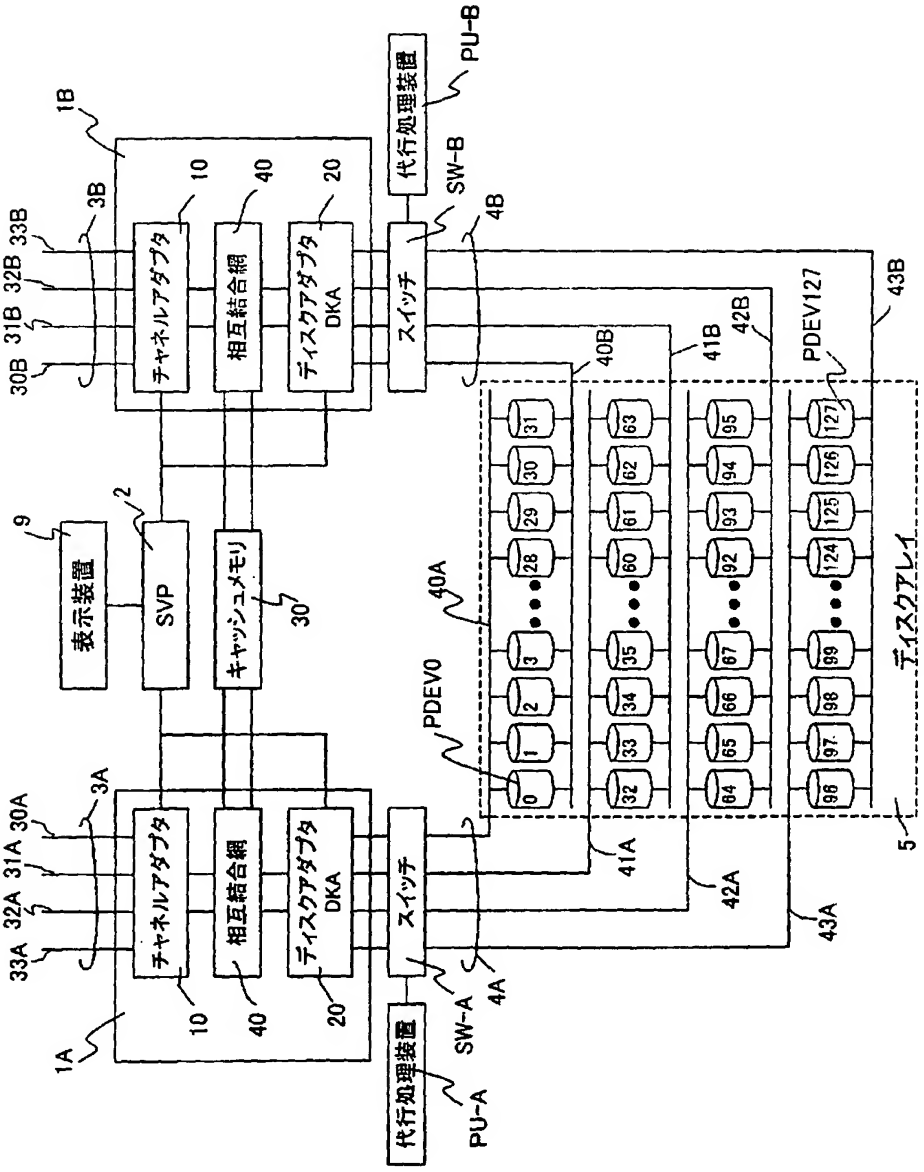
【図１７】

図　17


データ複写
400
SADDに複写開始アドレスを設定
401
EADDに複写終了アドレスを設定
402
BLに複写単位となる領域サイズを設定
403
CNT=SADD
404
405
CNT＜EADD
NO
終
YES
B
406
R/W命令あり
YES
NO
410
命令種別
READ
WRITE
411
連累ドライブからREAD
CNT～CNT＋BLのデータを連累ドライブから交替ドライブに複写
407
CNT=CNT+BL
408
412
複写済み領域へのアクセス
NO
YES
交替ドライブを含む新RAIDグループに対してWRITE
413
連累ドライブを含むRAIDグループに対してWRITE
414
A

【図１８】

図　１８


1B
3B
33B
32B
31B
30B
チャネルアダプタ
10
相互結合網
40
ディスクアダプタ
DKA
20
スイッチ
SW-B
4B
代行処理装置
PU-B
9
表示装置
2
SVP
キャッシュメモリ
30
40A
PDEV0
40B
41B
42B
PDEV127
43B
ディスクアレイ
5
0
1
2
3
28
29
30
31
32
33
34
35
60
61
62
63
64
65
66
67
92
93
94
95
96
97
98
99
124
125
126
127
41A
42A
43A
1A
3A
33A
32A
31A
30A
SW-A
4A
PU-A

【書類名】要約書

【要約】

【課題】　筐体サイズの小型化と記憶容量の大容量化に適し、ドライブ障害が発生しても、速やかに正常状態に復帰できるディスクアレイ装置および障害回復制御方法を提供する。

【解決手段】　上位装置と保守端末２に接続されたディスクコントローラ１と、ディスクチャネルを介して上記ディスクコントローラに接続されたディスクアレイ５とからなるディスクアレイ装置において、ディスクアレイにドライブ障害が発生した時、上記ディスクコントローラが、障害ドライブを搭載した障害基板上の複数のディスクドライブについて、交替ディスクドライブへのデータの移し替えを行い、論理グループの再編成が完了した後に、保守端末に障害基板が交換可能となったことを通知する。

【選択図】　図４

特願２００３－０６１４０３

出願人履歴情報

識別番号　［０００００５１０８］

１．変更年月日　１９９０年　８月３１日
［変更理由］　新規登録
住　所　東京都千代田区神田駿河台４丁目６番地
氏　名　株式会社日立製作所